GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 24, 2003, 11:50:07 ; Search time 1487 Seconds
(without alignments)
1025.[illegible] Mcells/sec

Title: US-09-463-480-9
Perfect score: 945
Sequence: 1 gnaggqtqttqqaattagqt.....tctcttctatatgttcaca 915

Scoring table: IGENITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154996 seqs, 8097714476 residues

Total number of hits satisfying chosen parameters: 32308792

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	59.4	6.3	1101	17	CNS0161A	AL107176 Drosophil
2	59.2	6.3	1101	17	CNS00L12	AL078714 Drosophil
c 3	57.2	6.1	1101	17	CNS00FVL	AL069706 Drosophil
c 4	55.4	5.9	899	17	AZ678725	AZ678725 ENTJF711F
c 5	55	5.8	734	17	CNS010MP	AL099163 Drosophil
c 6	54	5.7	979	17	CNS0161W	AL106190 Drosophil
c 7	53.2	5.6	1101	17	CNS0139G	AL063921 Drosophil
c 8	53	5.6	919	17	CNS0[illegible]RL	AL[illegible] Drosophil
9	52.4	5.5	1101	17	CNS0161L	AL106096 Drosophil
c 10	52	5.5	1101	17	CNS00A8G	AL064691 Drosophil
11	51.8	5.5	1038	17	CNS0108N	AL098657 Drosophil
c 12	51.6	5.5	928	17	CNS00D0KY	AL071865 Drosophil
13	51.4	5.4	1029	17	CNS00I7V	AL074642 Drosophil
14	51	5.4	1101	17	CNS00ZAL	AL097431 Drosophil
15	50.8	5.4	694	17	AQ853360	AQ853360 nbxb0000c
16	[illegible]	[illegible]	[illegible]	17	[illegible]	AL096962 Drosophil
17	50.6	5.4	981	17	CNS02B08	AL118970[illegible] Tetraodon
c 18	50.4	5.4	1065	17	CNS074RI	AL429184 clone BAC
19	50	5.4	871	17	AZ542905	AZ542905 ENTGA90TF
c 20	50	5.4	977	17	BH147612	BH147612 ENTGC47IF
21	50	5.4	917	17	AZ546499	AZ546499 ENTGS45IF
22	50	5.4	1101	17	CNS00EIW	AL069847 Drosophil
c 23	50	5.4	1289	17	CNS0160O	AL106578 Drosophil
c 24	49.8	5.4	946	18	BI911499	BI911499 604084[illegible]
c 25	49.8	5.4	956	17	CNS00F0H	AL071063 Drosophil
c 26	49.8	5.4	1024	17	CNS005KI	AL058090 Drosophil
c 27	49.8	5.4	1084	17	CNS00255	AL097258 Drosophil
c 28	49.6	5.2	1094	17	CNS012FZ	AL101513 Drosophil
29	49.2	5.2	1102	17	CNS014F4	AL104434 Drosophil
30	49	5.2	945	17	CNS006WN	AL066038 Drosophil
c 31	49	5.2	994	17	CNS04N0J	AL298972 Tetraodon
32	49	5.2	1037	17	CNS004M9	AL052822 Drosophil
33	49	5.2	1101	17	CNS00E07	AL069440 Drosophil
c 34	49	5.2	1184	17	CNS04P4P	AL300850 Tetraodon
c 35	48.6	5.1	671	17	CNS04P59	AL254213 Tetraodon
c 36	48.6	5.1	881	17	CNS06CAK	AL392274 18 end of
37	48.6	5.1	1015	17	CNS00J7I	AL075824 Drosophil
c 38	48.4	5.1	759	17	CNS06QXV	AL411257 17 end of
39	48.4	5.1	859	17	AZ674588	AZ674588 ENTJW161F
c 40	48.4	5.1	1201	17	CNS01613	AL106161 Drosophil
41	48.4	5.1	1203	17	CNS015WD	AL106008 Drosophil
42	48.2	5.1	1001	17	CNS01400	AL103554 Drosophil
c 43	48.2	5.1	1101	17	CNS00LJI	AL068807 Drosophil
c 44	48.2	5.1	1101	17	CNS00EP0	AL069498 Drosophil
45	48.2	5.1	1201	17	CNS0167M	AL106496 Drosophil

ALIGNMENTS

RESULT 1
CNS0161A
LOCUS CNS0161A 1101 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence 17 end of BAC BACN16N18 of BrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL107176
VERSION AL107176.1 GI:5626000
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billault at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRT project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.
FEATURES Location/Qualifiers

```
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /strain="BACR48P19"
                     /clone_lib="BACR48P19"
                     /note="end : 17"
BASE COUNT      156 a    129 c    184 g    293 t    339 others
ORIGIN

  Query Match             26.8%;  Score 594.4;  DB 17;  Length 1101;
  Best Local Similarity   28.1%;  Prob. No. 6.000e-29;
  Matches   96;  Conservative  100;  Mismatches  146;  Indels    9;  Gaps    9;

Qy       72 ATAGTGTATGATTATTAGTGATTCAATTATTTATAATTTCTCTTCTTGTATGTTCAAAT 131
           ::::   ::   : :|  |:  :||   ::  | |: || :  :: : : :|| :   :|
Db      720 RRAKsYGYTGTAYTTKTYTRTTTWATAAMWTTTATRTTTVVMGTTGKDRBAVTGTTGGTMT 785

Qy      132 AACTGGTCTTTAAATGAATAGATAATTAAGTTTGTAGCAATTTCTTCCTCAAATTGAGT 191
           : |  :: | ::   :|| |||      :|||   :|:|| | |:    : | :: |
Db      786 TKKTMRRRTTWCWRTKTCTCTATAMMMTRTTTTTAGAAATATATWTWTTsKsAYRsTTT 845

Qy      192 ATCAAAAA.TTTAAAATTTTTTTTTATAATATTTGATATAATTGTTTTTAAAAATTTT 251
           : ::::: | :     : | ||||:     :|:|: :: : :| |:|:   :: : :
Db      846 TTTPVVRTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 905

Qy      252 TATTTGAAAAGATTGTTATTATTCATTATGTTGTTGGACAATTGTTAGAGCCCCATCGCT 311
           |: :::: |:::|:    :::: :|:  |:|  :: : :::::| :     :: : |
Db      906 TWTWSRATARTWTRTTAAWAMMMMTTTATTTRTRTTTTRTTTTRTTTTRTTTTTTTT 965

Qy      312 AATGTCTTATATTATCTGAAATATTTTGGTAATAGAAGATAAAAAATCTCATTCTTTGTA 371
           ::: :::::|| :::::: : : :: : : : : : : : : : : : : : : : :
Db      966 TRTRRTWTATWTTTTWARKKKTAWTWATRATAWTTATTTRRTTTTTTTTTTTTTTTTT 1025

Qy      372 GTAGGAGTAAAAATTTGTCTTTTCATTATTCCATTGAATGTT 413
           ::: | :|    :|| | : :|   |: :   ||  |::
Db     1026 TTWTCRRTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1067


RESULT 2
CNS*VLIZ
LOCUS       CNS*VLIZ                1292 bp    DNA     linear   GSS 14-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence IEI3 end of BAC:
            BACR48P19 of BPCI 98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL078714
VERSION     AL078714.1  GI:5102004
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see: http://www.fruitfly.org/sequence/human-finished.html
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named BPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            PI and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila.htm.
FEATURES    Location/Qualifiers
     source          1..1292
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR48P19"
                     /clone_lib="BPCI 98"
                     /note="end : IEI3"
BASE COUNT      469 a    ?  c    ?  g    151 t    4 others
ORIGIN

  Query Match             6.3%;  Score 99.2;  DB 17;  Length 1101;
  Best Local Similarity   21.1%;  Prob. No. 0.00033;
  Matches  120;  Conservative  193;  Mismatches  256;  Indels    1;  Gaps    1;

Qy       41 TTTAGAGAAATATTTAAAATTTTTTATTTATAGTTATAGTTTAGTTTAAATTA 100
           ||||:     |: | :|::::: :::| | ||   :| : ::| |: :   || ::|
Db      495 TTTWRKATTTTWTTTTTAATTKWAAWWTTTTTAWTTWTTWATTAATATTTTTTTTKWTT 554

Qy      101 TCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAA 160
           : | |::|:: :    |  :|   |    |||  : ||:|| |: :: :::|:|||
Db      555 WKATRTTWTKKAATAATAATAAWAAWWTTTTTTAATWTTTTTWTTAKRTWAWRWWAWTTAT 614

Qy      161 GTTTTGTAGTAATT: :TCTTCAAATTTGATTATTAATGTTTTTTTTTTTTTTTTTTTT 220
           |:|| :   :|:: :|  :|   ||:     |: :|: :|  :|   |:|:|:|||
Db      615 TTWTTWTTTTWTTTTTWAATTTTWTTWTTAWATTTTTTTTTTTTAATKTTTTTTTTTTAT 674

Qy      221 TATATTTGATATAATTGTTTTGTAAGAATTGTTATTGAAAAGATTGTTTATTTTATTTTT 280
           || ||   | |: :: :   :: :::::|  : : :|:|: :|:|:|:|:|:|:|:|
Db      675 TAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 734

Qy      281 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 340
           : :: : ::: | |      : :  :: :  : :: :  :::|:|:|:|:|
Db      735 AAWAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 794

Qy      341 GAATAGAAGATGAAAAATCTTATTCTTTGTAGTAGGAGTAAAAATTTGTCTTTTCATTAT 400
           : ::   ::    :  : : :: :: : ::: |||| :::| :: : : : :
Db      795 KTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 854

Qy      401 TCCATTGAATGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 460
           :  ::: ::::      :  :: | :: :::| :| : : :  :  :  :  :  :
Db      855 KTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 914

Qy      461 AGTG-ATCTTTGGGATATTGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 519
           : :|  :  ||||| : ::| :: :: |:::   |:   :::::::::| :: :  :  :
Db      915 KTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 974

Qy      520 GAGAGAAAAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 579
           : |:: :|   : ::: :  :  : :::| |    :: :: :: :::   :: :::
Db      975 AAAAAWTTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1034

Qy      580 GATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 639
           : :::  :  ::: :|   : :|:|: :|:|
Db     1035 DATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1094


RESULT 3
CNS*EVIZ
LOCUS       CNS*EVI                 1101 bp    DNA     linear   GSS 04-JUN-2000
DEFINITION  Drosophila melanogaster genome survey sequence 2" end of BAC:
            BACR42B23 of BPCI 98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION     AL069706.1  GI:4949849
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (02 JUN 1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : sequ@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see: http://www.fruitfly.org/ The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI 98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            PI and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://[illegible]
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR39P23"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
BASE COUNT      419 a     91 c     60 g    299 t     232 others
ORIGIN

  Query Match             6.1%;  Score 57.2;  DB 17;  Length 1101;
  Best Local Similarity   36.4%;  Pred. No. 0.001;
  Matches  136;  Conservative   62;  Mismatches  175;  Indels    1;  Gaps    1;

Qy      37 TAGGTTTAGACAAATAGTCAAAATTGTCCTATTCTATAGGCATGATTTACTAGTCAGTTA 96
           ||  :|:: | :|||:   |:|:||   |:  |||:   :|  | |  |  |   ||
Db     977 TAWAWTWWAATWAATWATAWAWTTTAATAWWAATATWAWWWTWTTATTTTTTTTTTTT 918

Qy      97 ATTATTTTATAATTTCTTTTTTTTTTAT[illegible]TTTTTAATTAATTTATAA 156
           ||| :  |:|: :|    :: :: :   :  :|:|:  |  || :|| : :  :| || |
Db     917 WTTTWTATWTWTTTATATWWTWWAAAAWAWWAAWTWTATTTTTTWTTATWATTWATATTA 858

Qy     157 TTAAGTTTTGTAGCAATTTCTTCCTCAAATTGACTATCAACAATTGTTAGATTGCTTTGC 216
           :   |||: :   ::|  |  | : ||: : : | :| :  | |     |  :|:
Db     857 WATTTTTTWTWTATTWWTATATWATWWTATWTWTTTWATWTATWTATTTATWTWTWTT 798

Qy     217 TGATTATATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCAT 276
           : |||:|||:   ::: :: |:: ||: :  | || : | |  :| : | |  |  |
Db     797 WTATTWTATWTATAWWWTWWATWWTAWAAWTATATAWWATATATWWTWATTWTTWTT 738

Qy     277 TTCGTTGTTGGATGAATTGTTAGAGTCCCATCGCTAATGTCTTATAGTACTCGAAATATG 336
           ||  |: ||  |  ||: |:          ||  |  |  :|| ||
Db     737 TTATTWTTTTAAWTTATWTATWATTWTTWATWTTTTTTTTTTTTTWTATTTWAA-ATT 679

Qy     337 TTGGGAATAGAAGATGAAAAATCCCATTCTTTGTAGTAGGAGTAAAAATTTGTCTTTTCA 396
           |   | ||  :: |: |:  ||   ||  :|  |   |  |||||||||   :   |
Db     678 ATTTAATTAWWWAATWAWTTATTTAATAAWTAATTAAATTATTAAAAATTTTATWAAAAA 619

Qy     397 TTATTCCATTGAAT 410
           |:||   ||  ||
Db     618 ATATTTTTTTATAT 605


RESULT 4
AZ678725/c
LOCUS       AZ678725                 899 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTHP7TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic DNA sequence.
ACCESSION   AZ678725
VERSION     AZ678725.1  GI:11815871
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 899)
  AUTHORS   Lottus,B.J., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Lottus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 0543
            Email: bjlott@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13 Forward
            Class: shotgun
            High quality sequence start: 38
            High quality sequence stop: 855.
FEATURES             Location/Qualifiers
     source          1..899
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pBC1; Site_1: BstX I; Construction at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S. (1991) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:459.). The DNA was mechanically sheared to give a
                     tight size distribution ( 2 kb). The y + i method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vandin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      [illegible]
ORIGIN

  Query Match             5.9%;  Score 55.4;  DB 17;  Length 899;
  Best Local Similarity   55.4%;  Pred. No. 0.0028;
  Matches  107;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy      67 ATTCTATAGGCATGATTTACTAGTCAGTTAAATTATTCTATAATTTTTTTTTTTTTATAT 126
           |||||  |  |||  ||:|| || ||| |||||||    :|   ||    |    ,  |
Db     478 ATTCTTTTAGCAGTATTTATTAATCATTTAATTAAATGCATTTCTTTCAGTTAAAATTATGCA 419

Qy     127 CAAATAATTTTTTCTTTAATTAATAATATAATTAATTTTTTTTATTTTTTTCAAAT 186
           ||  ||| |  |  |||      | |||||  ||  || |  |  :|    || | ||
Db     418 GCAACAACCAGAGCAATAATATCATCAAAATTAACATTTTTAACTATCATATCATAAATT 359

Qy     187 TGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTGATATAATTTGTTTCTAACA 246
           ||| || ||   ||  |  |  || ||| | || ||| |  || ||||  ||  :| || ||
Db     358 TGATTATAAAATTTTTCTGTCTGTGGTGTCATGTTTAAAATTAATCATTTTCCTTTTAGTA 299

Qy     247 ATGTGTAGTGAAA 259
           || | :| |  |:
Db     298 ATATTTACTTCAA 286


RESULT 5
CNS01OMP/c
LOCUS       CNS01OMP                 734 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN041Z9 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AI699163
VERSION     AI699163.1  GI:5616774
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
```

```
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 734)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP)
            http://www.bdgp.fsm.edu.uk . This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..734
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN41P7"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
BASE COUNT      298 a     62 c     24 g    211 t     171 others
ORIGIN

  Query Match             45.8%;  Score 55;  DB 17;  Length 734;
  Best Local Similarity   88.8%;  Prob. No. 0.0004;
  Matches  143;  Conservative   59;  Mismatches  162;  Indels    5;  Gaps    2;

Qy      41 TTTACAGAAATACTCAAAATTGTGTATTCTATAGGCATCATTACTACTCAGTTAATT 99
           |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     445 TATAAACAAWTAATAAWAATATTGTCWTTTTATGTTGTTATAAWWTATKWATWTGTA 376

Qy     100 ATCCTATAATTTCTCTCTTGTATGCTCAAATAACTGTTCCTTTAATGAATAGATAATTA 159
           || |:| |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|||
Db     375 ATAKTWTTTATRTGKTTWTTTWTATTTAWTWWTTWTATTTWTTTTTWWTTKATAWTTT 316

Qy     160 ATTTTTTTA-AA---TT-T-AAT---ATTAT-AATAATTTTAAAAAAAAAAAAAAAAA 214
           |:|:|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     315 TGTATTWTTTTTCTKWWAAATTTAATATWCWTTTTKATTTTTTGTTCWWTTTTWRTATAW 256

Qy     220 TTATATTTGATATAATTTTTTTTAATAATTTTTATTGAAAATATTTGATTATTTCATTTC 279
           |||||||:|:|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     255 WWWWWKTTTTTTWWATCATTRTWTT    TTTAAWTWTAATTATAAWTTWTTTTWTTTWA 200

Qy     280 GTTGTTGGWCAATTGTTAGAWCTCCATCGCTAATCCTTATATGTACTCGAAATATGTTG 339
           :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     199 WKWTTWATWAWAAAAATWTRCTTWWWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 140

Qy     340 CGAAAAAAAACAAAAAA-CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 398
           :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     139 AAAAAAAAAAAAWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 80

Qy     400 TTTCATTGA 408
           |:|:|:|:|
Db      79 ATWATTWTA 71


RESULT 6
CNS0190_W/c
LOCUS       CNS0190W                  970 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN16B17 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL100190
VERSION     AL100190.1  GI:5626078
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP)
            http://www.bdgp.fsm.edu.uk . This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..970
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN16B17"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      275 a     79 c     117 g    317 t     182 others
ORIGIN

  Query Match             45.7%;  Score 54;  DB 17;  Length 970;
  Best Local Similarity   96.1%;  Prob. No. 0.0006;
  Matches  100;  Conservative   61;  Mismatches  113;  Indels    3;  Gaps    1;

Qy      82 TTTAGTAGTGAGTTAATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGTTCCT 141
           ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     979 TTTTTTTAWAATTWTTTTTTTHHHBTATTWTCTTTTTTTTTTTATTTTTWTNTTTTTTATTB 920

Qy     142 TTAATGAATAGATAATTAAGTTTTGTAGCAATTTCTTCTCCTCAAATTGAGTATCAACAATT 201
           |:||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     919 TAWATAAAAAAAAAAAAAATAWAWWTWWTTWTAWTTTTAWATWBAATTWWTHTTTTHBBA 860

Qy     202 GTTAGATTGTTTGGTGATTATATTTGATATAATTTGTTTGTAAGAATTGTGTATGAAAAG 261
           ||  |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     859 TTT  NTNWWATYTTMTWTTTTTWTWATWTMTW...TTT.AWAAAAAAAAAAAAAAAAAAAAA 803

Qy     262 ATTTTGATTATTTATTTTTTTTTTTTTTGATGAATTGTTAGATTTTTATTTTTTTTTTTTT 321
           ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     802 ATTATMTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 743

Qy     322 AGTATCTGAAATATGTTGTTGGWCAATTWGWATTWAWTTTTTTTTTTTTTTTTTTTTTT 381
           |:||:|:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     742 AKTAWTTWAAAATTWTTTTTRWTTWAWWWTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 683


RESULT 7
CNS0003G3/c
LOCUS       CNS0003G3                 1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
```

melanogaster genome using these BACs. For further information
please see: http://www.fruitfly.org/p_info
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BkmF's
PI and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.chori.org/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_lib "RPCI-98"
/note="end : IE13"

BASE COUNT 201 a 64 c 131 g 282 t 503 others
ORIGIN

Query Match 5.6%; Score 53.2; DB 17; Length 1101;
Best Local Similarity 12.2%; Pred. No. 0.011;
Matches 62; Conservative 157; Mismatches 103; Indels 5; Gaps 3;

Qy 131 TAACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTTCTTCCTCAAATTGAG 190
Db 1100 KARRWCNSSWWSRTRRSTWWTRWWSWWSBACSSSWACSASSWAWSSCACSWSATWWW 1041

Qy 191 TATCAACAATTGTTAGATTGCTTTGGTGATTATATTTGATATAATTGTTTGTAAGAATGT 250
Db 1040 WWWWWATWDTWWDFWWWWATAAFSSCAWTWWSTAWSATWAGSSCASSSDADAATDADSA 981

Qy 251 GTAGTGAAAAGATTGTGATTATTCATTTCGTTGTTGGACGAATTGTTAGAGCCCCATCCC 310
Db 980 GRRDGGRFFSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 921

Qy 311 TAATGCCTCATAGTACTCGAAATATGTTGGGAATAGAAGATGAAAAATCCCATTCTTTGT 370
Db 920 DDDDGAGDSTTSSKGKLATSSTTDSTSKPDDESFFWSWSWSFAASTWSDATAWWWWG 861

Qy 371 AGTAGGAGTAAAAATTTGTCTTTCATTATTCCATTGAATGTTAACACTTGCCATTCAT 430
Db 860 WADASWTSWSAAATSWWASSWSWSAWAWSWSSSWSASATALFFWSSSASSSSARRRRD 801

Qy 431 CTGACGGGGATGGCAGAGTTCCGACCATCTAGTGATCCGTGGGATATTGATTTTGGTGTG 490
Db 800 RRPADDRPFAADDPDDAATWTTWTTTTTPDTDDWPWPTTPWTPWAADPTWDRDDDDDRDR 741

Qy 491 TCAATGAAATTGTGAGAACGGGCTTCTGGGAGAGAAAAGCCCTCTTGCCTCTGATATGAA 550
Db 740 AJTACFFWFSTWFFFWFFSTTFWSSASFSCAFSSSSSSSSSSSSSSSSSSSSSSSSSS 681

Qy 551 CACTGAGGCTGATTATGTTAACGGATGGAGATTTATCAGTGGCTGAATTTGGGTGGTGTA 610
Db 680 RATWDRTDAWWADAAWWTTTDTDTDDWDKRDRRRCSARRRRTTARAAWTAWTWKAWDWA 621

Qy 611 GAGACAGAATTTGAAAGTTCTAACAATAAA 640
Db 620 KWDWKTRATPWDKWAADTWTDARKADRDWA 591

RESULT 8
CNS005KL:
LOCUS CNS005KL 919 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence 17 end of BAC #
BACR12F23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL061409
VERSION AL061409.1 GI:4944512
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 919)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC end sequence was carried out as part of a
collaborative project to sequence the Drosophila Genome Project (DGP).
The DGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see: http://www.fruitfly.org/p_info
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BkmF's
PI and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.chori.org/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1..919
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR12F23"
/clone_lib "RPCI-98"
/note="end : 17"

BASE COUNT 273 a 124 c 99 g 197 t 226 others
ORIGIN

Query Match 5.4%; Score 51; DB 17; Length 919;
Best Local Similarity 25.6%; Pred. No. 0.011;
Matches 114; Conservative 117; Mismatches 192; Indels 2; Gaps 1;

Qy 9 TTGGAATTAGGTTTGCCTAGGGTTTGCCTAGGTTTAGAGAAATAGTTCAAAATTGTCCTAT 68
Db 891 KSSSSWSWSAWSWSKKKKKKSSSKSSKKSASAAAKAAAAKKSAASSWSWSSASSASSSK 832

Qy 69 TCTATAAGCATGATTTATTAGTAGTGAGTTAATTATCCTATAATTTCT CTTCTTGTATCCCT 126
Db 831 KSKSSSSWSTAWTSSWSWSWSWSSTATSAWSWASSSWSTSSSKSSSSSSAAATAAKSSW 772

Qy 127 CAAATAACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTTCTTCCTCAAAT 186
Db 771 KSWSSWAAWSWSWSWSSAAAWSSAASSTWTTTAASTKKKKSSSSSSSWAAASWSTATKAKWSWSK 712

Qy 187 TGAGTATCAATAATTGTTAGATTGCTTTGGTGATTATATTTGATATAATTGTTTGTAAGA 246
Db 711 KKKKKSSSWKWKWSTTWSSNAWWSWTTTTTTTTTTTTTWTTTTTTSSAAAWWTTTTSTTT 652

Qy 247 ATGTGTAGTGAAAAGATTGTGATTATTCATTTCGTTGTTGGACGAATTGTTAGAGCCCCA 306
Db 651 TTATATGTSSAAGTAWWSSSAAAA.TSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 592

Qy 307 TCCCTAAATGCCTCATAGTACTCGAAATATGTTGGGAATAGAAGATGAAAAATCCCATTCT 366
Db 591 TKKKSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 532

Qy 367 TTGTAGTAGGAGTAAAAATTTGTCTTTCATTATTCCATTGAATGTTAACACTTGCCATCT 426
Db 531 DSWWTTTTTTSSGSAASAATATTTTRSTWTTASWTATTTATTSNNSNTTSTTTTNSTTW 472

Qy 427 TCATC 431
Db 471 TTAAC 467

RESULT 9

```
CNS01GLI
LOCUS       CNS01GLI                1161 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106896
VERSION     AL106896.1  GI:5624374
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1161)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk/. The Drosophila melanogaster BAC
            library (DrosBAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1161
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN16D22"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
BASE COUNT     203 a    223 c    84 g    198 t    436 others
ORIGIN

  Query Match             5.9%;  Score 52.4;  DB 17;  Length 1161;
  Best Local Similarity   36.6%;  Pred. No. 0.017;
  Matches  104;  Conservative   87;  Mismatches  148;  Indels    1;  Gaps    1;

Qy         51 TACTCAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGACTTAATTATCCTATAATT 110
Db        736 WWAWAAAAAAATKAHDHHHHHHHHWHHWHHHHBHHHATWHHHAKDHHHHHH 795

Qy        111 TCTTCATTCTATACTCAAATAATCTGTTCTTAATGAATAGATAATTAAGTTTTGTAG 169
Db        796 WATATTTWTTWTTDATWTTTTTTTTTWAWTATATATTTWAWTATAATTAWAAATWWTATTA 855

Qy        170 CAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTGA 229
Db        856 WTATAWGATAWTWGWWWTKTAWWWTTWTWWWWATTAKTTWATWATATTAWTWAWAWWTAW 915

Qy        230 TATAATTTTTTTAAGAATTTTATTTAAAAAATTTTATTTTTTTATTTCTTTGTTGGAC 289
Db        916 AWWAWAATTATWTWTWATTATTWKTTTASTATWTTTTWTWATAWKAKAWKWATWLWAWA 975

Qy        290 GAATTGTTAGAGCCCCATCGCTAATGCCTTATAGTACTCGAAATATGTTGGGAATAGAAG 349
Db        976 WKATAWKTATWAATAKATTWATATAAAWWTWWTATTGWKWAWTAKATTAWTAAATKTAA 1035

Qy        350 ATGAAAAATCCCATTCTTTGTAGTAGGAGTAAAAATTTGT 389
Db       1036 ATWTWTTKWTTKWKTTKWTTKTAKTKAKAKRAWWWWRT 1075


RESULT 10
CNS003BD/c
LOCUS       CNS003BD                1161 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K08 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL064091
VERSION     AL064091.1  GI:4941847
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1161)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org. The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1..1161
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR08K08"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT     305 a    127 c    108 g    484 t    145 others
ORIGIN

  Query Match             5.5%;  Score 52;  DB 17;  Length 1161;
  Best Local Similarity   30.1%;  Pred. No. 0.021;
  Matches  143;  Conservative   48;  Mismatches  179;  Indels    5;  Gaps    1;

Qy         56 AAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGACTTAATTATCCTATAATTTCTCT 115
Db       1085 ATAKWWKKTKWKTWWTTWATAAWNTAKTTKTTTTTTKWTATTWATAKAWTTTTWTWTTT 1026

Qy        116 TCTTCATTCTCAAATAATCTGTTCTTAATGAATAGATAATTAAGTTTTGTAGCAATTT 175
Db       1025 WTTTRWWTWTWWATTTWATTTTATTTTTTTTTTTTRWWTATTATAATWATWATTTAAAAATA 966

Qy        176 CTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTGATATAAT 235
Db        965 TTWTWTWTAATWTWTTTWTAKAKAAAWAWTAWTWATTTTATWTTATTTTTTTAWTWT 906

Qy        236 TTTTTTAAGAATTTTATTTAAAAAATTTTATTTTTTTTATTTCTTTGTTGGACGAATTG 295
Db        905 TAWTTTAWTTATTTWAAAAWWTATWAWATAWATTATTTTATTTATTTTTTTWTTTTTTW 846

Qy        296 TTAGAGCCCCATCGCTAATGCCTTATAGTACTCGAAATATGTTGGGAATAGAAGATGAAA 355
Db        845 TANAAIWTTTAATKAWATTTAAWTATWAAATTTAAAWAAAWT     ATWNWNATATATA 791

Qy        356 AATCCCATTCTTTGTAGTAGGAGTAAAAATTTGTTTTTATTATTGCATCGAATGTTAA 415
Db        790 AATWATWTATTTTTAATATAAATAAAWTATWTTTTTTTAATATTTTWTAAAWAWTAT 731

Qy        416 CCACTTGCCATTCAT 430
Db        730 AATTTAWAWTTTAT 716


RESULT 11
```

```
CNS01DRN
LOCUS       CNS01DRN                1038 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN03L01 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AI098657
VERSION     AI098657.1  GI:5610268
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1038)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage -
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.[illegible] The DrosBAC
            library (Dros BAC) was made by Alain Billault at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1038
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN03L01"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      372 a      6 c    119 g    219 t    322 others
ORIGIN

  Query Match             5.5%;  Score 51.8;  DB 17;  Length 1038;
  Best Local Similarity   26.1%;  Prod. No. 0.024;
  Matches  151;  Conservative  147;  Mismatches  278;  Indels    3;  Gaps    1;

Qy      38 AGGTTTAGAGAAATAGTCAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTAA 97
           ||:::.  ||:||  | |:::    |   :|| :| |     |  | |     ||
Db     437 AGAEEAAAAAAAAATAAAAA.ADWDAAAAAAAA.WLTGP3kAAAsAAAAAAAA.A.?AAAA 496

Qy      98 TTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAAT 157
           |   ||:  ||| |       ||  : ||| || |  :  : ||  || | | |
Db     497 AAAAAATAETTTTTATAAAAAAAAATATPAAAAAAAAA?TTPAAPAAAAAAAAAAAAA?A 556

Qy     158 TAAGTTTTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGT 217
           | ||:: :  ||   ::   |       : |  :|| ||   |  :|   :|: ::: :: :|
Db     557 TDWTTTTLAAAAAPGPAAA?.AAPAA.AAAAAtAlWtPA?L.APWtAPWtAPtA 616

Qy     218 GATTATATTTGATATAATTGTTTGTAAGAATGTTAGTGAAAATATTTTAATTATTCATT 277
           |:||  :::|||:::  :::  :  :  :: : : :|  :|| :: :: :|
Db     617 ATAAAAWATTAAWPATATTTT.AAPPA.L.PPAAA.L.P?AAWLLPPtLAPWLPPAPA. 676

Qy     278 TCGTTGTTGGACGAATTGTTAGAGCCCCATCGCTAATGCCTTATAGTACTCGAAA--TA 334
           : ::::::: | |  :::| | |     :: : :  :|   :|  ::    :  |   :
Db     677 KPPPWPPPPATAPPPPPPPGAAAAPAAAPPAPPPPPPPPRAGCPPPPPPPPPPPPP 736

Qy     335 TGTTGGAATAGAAGATGAAAAATCCCATTCTTTGTAGTAGGACTAAAAATTTGTCTTTT 394
           |||:: :|::|  | : | ||||   |:   |||:|  | : ::: : |:::| | :|||:
Db     737 TGTPPAPAWWAWPATCPTA.AAAPAAAAAAWPPPPPPPPPPPPPPPPPPPPPPPPPPP 796

Qy     395 CATTATTGGATTCAATGTTAACCACTTGCCATTCATCTGACGGGGATGGCAGAGTTCCGA 454
           |  :|  :|  | :|    ||:  :  ::  :||: :|| |||   |
Db     797 KTTPTTAAWWTTTTTTPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 856

Qy     455 GTATCTAGTGATTGGTGGGATATGATTTGTTTGTCAATCAAATTTTGATAAACGGGCT 514
           :  : :::  |::: :::  ::|:||: ::|:  ::  ::|::       :
Db     857 AGGPCTPP?APPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 916

Qy     515 TGTTAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 574
           | : : |  ::| :  : : |   ::|:|  :  :  :  |  || |||| |
Db     917 PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 976

Qy     575 AGGAGAATTAATAGTGGTGAATTGGGTCGTGAGAT 613
           |:  : | | |  ::: ::: |:|:|| |  | ::|
Db     977 PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 1015


LOCUS       CNS01DKY                 928 bp    DNA     linear   GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence 17 end of BAC #
            BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AI071865
VERSION     AI071865.1  GI:4948170
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 928)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage -
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            [illegible] The RPCI Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            BsoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            PI and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://[illegible]/rpci98.ordering.html.
FEATURES             Location/Qualifiers
     source          1..928
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR27A24"
                     /clone_lib="RPCI-98"
                     /note="end : 17"
BASE COUNT      252 a      2 c     84 g    321 t    269 others
ORIGIN

  Query Match             5.5%;  Score 51.6;  DB 17;  Length 928;
  Best Local Similarity   31.7%;  Prod. No. 0.026;
  Matches  121;  Conservative   85;  Mismatches  171;  Indels    7;  Gaps    1;

Qy      44 AGAGAAATAGTCAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTAATTATCC 103
           | | |||::  ::|||:   :: :  : : :  :  | |:: : : :::::: |
Db     866 AAAAAAWWYTCMAAAMAMAMETMCVMMMMMMMMMMMMMMMMMMMMWWHHHHHHHHWTTTT 807

Qy     104 TATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAGTT 163
           | |: ||| | || || :::  : ::::  :  :| :::::: ::|   : : ::|  |
Db     806 TTTWTTTTTTTTTTTTTWWWWWWWWWWWTTTWWWWTWWWWWWWWWWWTNNNWNHCHWAWAT 747

Qy     164 TTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTAT 223
           :: ::   |||| |   | |:|| | |  :| ||  :|  ||  |:|  | :|:|
Db     746 WWWWWWATATTTTTTTWTTWWAATAAAAAAAAAAAAAAAAAAAAAAATTTTTWTTTAWTTWT 687

Qy     224 ATTTGATATAAATGTTTGTAAGAAATGTTAGTGAAAATATTTTAATTATTCATTTCGTTG 283
           |    | : ||| ||| | :   | | |:  | :| |:| : : :  || .
```

```
Db      686 AAAATTTTTTATTWTTTTTWTTTTWTTTWTWTATATAAWAWTAWAWAAWAWATATTATAIW 627

Qy      284 TGGATTAATTGTTATATTTTATTATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 343
            |   |  ||| ||        ||   |   | ||  | | |||||  |  ||
Db      626 AWATATTTWWTATTW       AWTATAAAATWWTAWWTTTAAWTAAWATAWAATTAAAA 574

Qy      344 TATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAT 403
            | | |  ||||   | | |  |  |  |  |||||| | | |||| || | |  |  |
Db      573 WTWAAAWAAAAAAAAAWAAAATAWTTTWTWAAAAAWWAAAAAAWATTWTTTTTTWWTMATAA 514

Qy      404 ATTGAATTTTAAATTATTGTTATTT 427
            |  | ||||||   |  | |  |
Db      513 AAAATTTAWTAWTATATTATWTT 490


RESULT 13
CNS0017V
LOCUS       CNS0017V                1 29 bp    DNA     linear   GSS 15-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence TE13 end of BAC
            BACR38K07 of RPCI-98 library from drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL074042
VERSION     AL074042.1  GI:4954220
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1029)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91 06 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using those BACs. For further information
            please see: http://www.fruitfly.org the BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain yz; cn bw sp, the same strain used for the BDGP's
            PI and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.chori.org/Drosophila_ke.htm.
FEATURES             Location/Qualifiers
     source          1..1029
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR38K07"
                     /clone_lib="RPCI-98"
                     /note="end : TE13"
BASE COUNT      267 a    95 c    179 g    161 t    337 others
ORIGIN

  Query Match          5.4%;  Score 51.4;  DB 17;  Length 1029;
  Best Local Similarity   72.7%;  Pred. No. 0.095;
  Matches   91;  Conservative   16;  Mismatches  188;  Indels    0;  Gaps    0;

Qy       1 GGAGGTTGTTTGGAATTAGTTTTTCTCTAGGTTTGCCTAGGTTTAGAGAAATAGTCAAAAT 60
           ||  || | ||  ||||| || |  | ||| | | ||| | | | | | | |  |
Db     550 GGKTCTKMCAACCACAGMCCKACTTCCTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 609

Qy      61 TGTCTTATTCTATAGTCATGATTTAGTAGTAGCATTTAATTATCTCTATAATTTCTCTTCTTG 120
           |  |  |  |  | | |  |  | |  |  |||  |  |  |  |  |  |  |  |||
Db     610 CAKTCMTCCTCTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 669

Qy     121 TATGTCAAATAATTTTTTCTTTAATGAATAGATAATTAAGTTTGTAGCAATTTCTICC 180


Db     670 TATTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 729

Qy     181 TAAATTTATTATTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 240
           |||  | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     730 KATWWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTA 789

Qy     241 GTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 300
           |||  | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     790 DDBTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 849

Qy     301 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 360
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     850 GTKTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 909

Qy     361 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 420
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     910 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 969

Qy     421 TTTTTTTTTTTTTTTTTTT 439
           | | | | | | | | | | |
Db     970 TTTTTTTTTTTTTTTTTTT 988


RESULT 14
CNS002AL
LOCUS       CNS002AL                1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN01120 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL097431
VERSION     AL097431.1  GI:5600642
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP)
            http://www.sdp.cbil.upenn.edu. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billault at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            p[illegible]. The DNA was prepared from embryos by Alain Bucheton
            and Geneviève Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN01120"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      301 a    216 c    197 g    149 t    138 others
ORIGIN

  Query Match          5.4%;  Score 51;  DB 17;  Length 1101;
  Best Local Similarity   84.7%;  Pred. No. 0.098;
  Matches   94;  Conservative   14;  Mismatches  117;  Indels    0;  Gaps    0;

Qy       1 GGAGGTTGTTTGGAATTAGTTTTTCTCTAGGTTTGCCTAGGTTTAGAGAAATAGTCAAAAT 60
           |||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db     935 GWWGGTATTTKKTKTCTCAATVTAVVVAMNTYTTTTTTTAWTTWAWAAAAWAWTAWTTT 994

Qy      61 TGTCTTATTCTATAGTCATGATTTAGTAGTAGCATTTAATTATCTCTATAATTTCTCTTCTTG 120
           |  |  |  |  | | |  |  | |  |  |||  |  |  |  |  |  |  |  |||
Db     995 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1054
```

```
Db    896 TTTTATWTAWTAAAAAAAAWWTTTTTTTTTTTTTTTWTTYTTTTTTTATWTATAWTWWWT 955

Qy    121 TATGCTCAAATAACTCGTTCTTTAATGAATAGATAATTAACTTTTGTAGCAATTTCTTCC 180
          : |  | :|:::| :  || :: ||: ||:| | ||  ||  :  ||  :|||  ||
Db    956 KTTTTTTWAHHWAAYTWTTTKWAAAWAAAWAAAAAAAAAAAAAWAAATAWAWATTRTTTTT 1015

Qy    181 TCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTGATATAATTGTTT 240
          | : ::  :  | |:   || ||   :   :|  | :|| | ||  ||:|| || |||
Db   1016 TTWTWWAMWWAATTAWTTTTTTTTTTTWATAAWTTTTTWTTTTTTTTTTTWTATTTTTTT 1075

Qy    241 GTAAGAATGTGTAGTGAAAAGA 262
            |: | | | |: | ::| |
Db   1076 AAAWAATTWTWTWTTTTWWAWA 1097


RESULT 15
AQ853360
ID   AQ853360        AQ853360.                694 bp    DNA     linear   GSS 03-NOV-2000
DEFINITION  nbxb0003G02r CUGI Rice BAC Library Oryza sativa genomic clone
            nbxb0003G02r, DNA sequence.
ACCESSION   AQ853360
VERSION     AQ853360.1  GI:6203817
KEYWORDS    GSS.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GCAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence stop: 143.
FEATURES             Location/Qualifiers
     source          1..694
                     /organism="Oryza sativa"
                     /strain="Japonica"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="nbxb0003G02r"
                     /clone_lib="CUGI Rice BAC Library"
                     /tissue_type="Leaf"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Rice is one of two most popular grains in the
                     world. Half of the world  population especially those
                     inhabiting highly populated areas of the  humid tropics
                     and subtropics, rely on rice as their primary source of
                     carbohydrate. Monocotyledonous rice is a diploid plant
                     (2n=24) with a  haploid genome equivalent of 441 Mbp
                     (Arumuganathan and Earle, 1991).  The relatively small
                     genome of rice, three times larger than that of
                     Arabidopsis, makes it suitable for genomic studies. In
                     order to  facilitate positional cloning, physical mapping
                     and genome sequencing  of rice, we have constructed a BAC
                     library from Oryza sativa, Nipponbare  variety. The
                     library contains 36,864 clones with an average insert size
                     of 128.5 Kb providing 10.9 haploid genome equivalents. The
                     deep  coverage allows the isolation a particular sequence
                     with a probability  of 99.9 %. Two high density filters,
                     each containing 18,432 clones  (doubly spotted), represent
                     the whole library for colony screening."
BASE COUNT      257 a     27 c     44 g     497 t      59 others
ORIGIN
```

```
Query Match             5.4%;  Score 56.8;  DB 17;  Length 694;
Best Local Similarity   51.8%;  Pred. No. 0.038;
Matches  113;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy     56 AAAATTGTGCTATTCTATAGGTATGATTTATTATAGTTGTTAATTATTCTATAATTTCTCT 115
          | | || |  ||| ||     ||   ||  |  | | || ||  |  |  | : || | |
Db    365 ATATTTTTATTATTTATTTATATATATTTATTTTTAAATTTATATTTTTAAATTTTTTTT 424

Qy    116 TCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTT 175
          | |  ||   |  |||| |  || ||| ||    |   || ||  || | :|     ||
Db    425 TATATTAATTTATTATAATTTTTTTTTTTTATTTTTTTTTTATTTATTTTATATATTTTATT 484

Qy    176 CTTCCTCAAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTGATATAAT 235
          || | ||||| | ||| ||  ||  ||  |  ||| | ||| ||| | || |
Db    485 ATTTTTCAAATTAATTATTAATTTTTTATATTATTTTTTTTTTATTTAATTTTTTTTTTATT 544

Qy    236 TGTTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCATTTCTTTGTT 285
          | | |     || | |  | | | |||     ||| |||  : ||
Db    545 TTTATTATTTTATTTTTTTTTATATAAATTTATTATATTTTTTTTTTATTTT 594


Search completed: February 24, 2003, 14:45:25
Job time : 1499 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 24, 2003, 10:11:36 ; Search time 2650 Seconds
(without alignments)
10378.179 Million cell updates/sec

Title: US-09-463-480-9
Perfect score: 945
Sequence: 1 ctaaaatttttaaattaaat.....tctcttctatcatttcaca 945

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2954640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Postprocessing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htq:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htq_hum:*
31: em_htq_inv:*
32: em_htq_other:*
33: em_htq_mus:*
34: em_htq_pln:*
35: em_htq_rod:*
36: em_htq_mam:*
37: em_htq_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	88.6	8.8	625	8	AF110779	AF110779 Lilium lo
2	56.2	5.9	27291	2	AC115575	AC115575 Dictyoste
3	55.8	5.9	25117	2	AC116102	AC116102 Dictyoste
4	54.6	5.8	7814	6	AX251284	AX251284 Sequence
5	54.2	5.7	294914	3	AE003603	AE003603 Drosophil
6	53	5.6	7851	6	AX346662	AX346662 Sequence
7	53	5.6	7851	6	AX346663	AX346663 Sequence
8	53	5.6	191440	9	AC012645	AC012645 Homo sapi
9	52.2	5.5	6673	6	AX458595	AX458595 Sequence
10	52.2	5.5	93475	5	AL672088	AL672088 Zebrafish
11	51.6	5.5	1600	3	AB048741	AB048741 Bacteroid
12	51.6	5.5	1600	3	AB048742	AB048742 Bacteroid
13	51.6	5.5	1602	3	AB048752	AB048752 Bacteroid
14	51.4	5.4	6281	3	AF465367	AF465367 Dictyoste
15	51.4	5.4	61052	2	AC117074	AC117074 Dictyoste
16	51.4	5.4	61052	2	AC124513	AC124513 Dictyoste
17	51.4	5.4	233479	2	AC126758	AC126758 Homo sapi
18	51	5.4	6126	6	AX251327	AX251327 Sequence
19	51	5.4	6126	6	AX346732	AX346732 Sequence
20	51	5.4	8093	6	AX251189	AX251189 Sequence
21	51	5.4	[illegible]	3	AX281166	AX281166 Sequence
22	51	5.4	8093	6	AX347418	AX347418 Sequence
23	51	5.4	[illegible]	6	AX345161	AX345161 Sequence
24	51	5.4	8093	6	AX456411	AX456411 Sequence
25	51	5.4	274172	2	AC122488	AC122488 Mus muscu
26	50.6	5.4	15732	6	AX277932	AX277932 Sequence
27	50.6	5.4	15732	6	AX324619	AX324619 Sequence
28	50.4	5.3	11790	6	AX345445	AX345445 Sequence
29	50.4	5.3	167098	5	AL590149	AL590149 Zebrafish
30	50.4	5.3	175159	2	AC097877	AC097877 Rattus no
31	50.4	5.3	228208	2	AL732510	AL732510 Danio rer
32	50.4	5.3	349980	6	AX344558	AX344558 Sequence
33	50.2	5.3	5984	6	AX458478	AX458478 Sequence
34	50.2	5.3	6061	6	AX345043	AX345043 Sequence
35	50.2	5.3	6668	6	AX251172	AX251172 Sequence
36	50.2	5.3	6669	6	AX344755	AX344755 Sequence
37	50.2	5.3	6668	6	AX346121	AX346121 Sequence
38	50	5.3	1141	6	AX083744	AX083744 Sequence
39	50	5.3	1600	3	AB048743	AB048743 Bacteroid
40	50	5.3	1600	3	AB048744	AB048744 Bacteroid
41	50	5.3	1600	3	AB074023	AB074023 Bacteroid
42	50	5.3	1601	3	AB035111	AB035111 Bacteroid
43	50	5.3	1611	3	AB035119	AB035119 Bacteroid
44	50	5.3	5070	6	AX345704	AX345704 Sequence
45	50	5.3	6286	6	AX251445	AX251445 Sequence

ALIGNMENTS

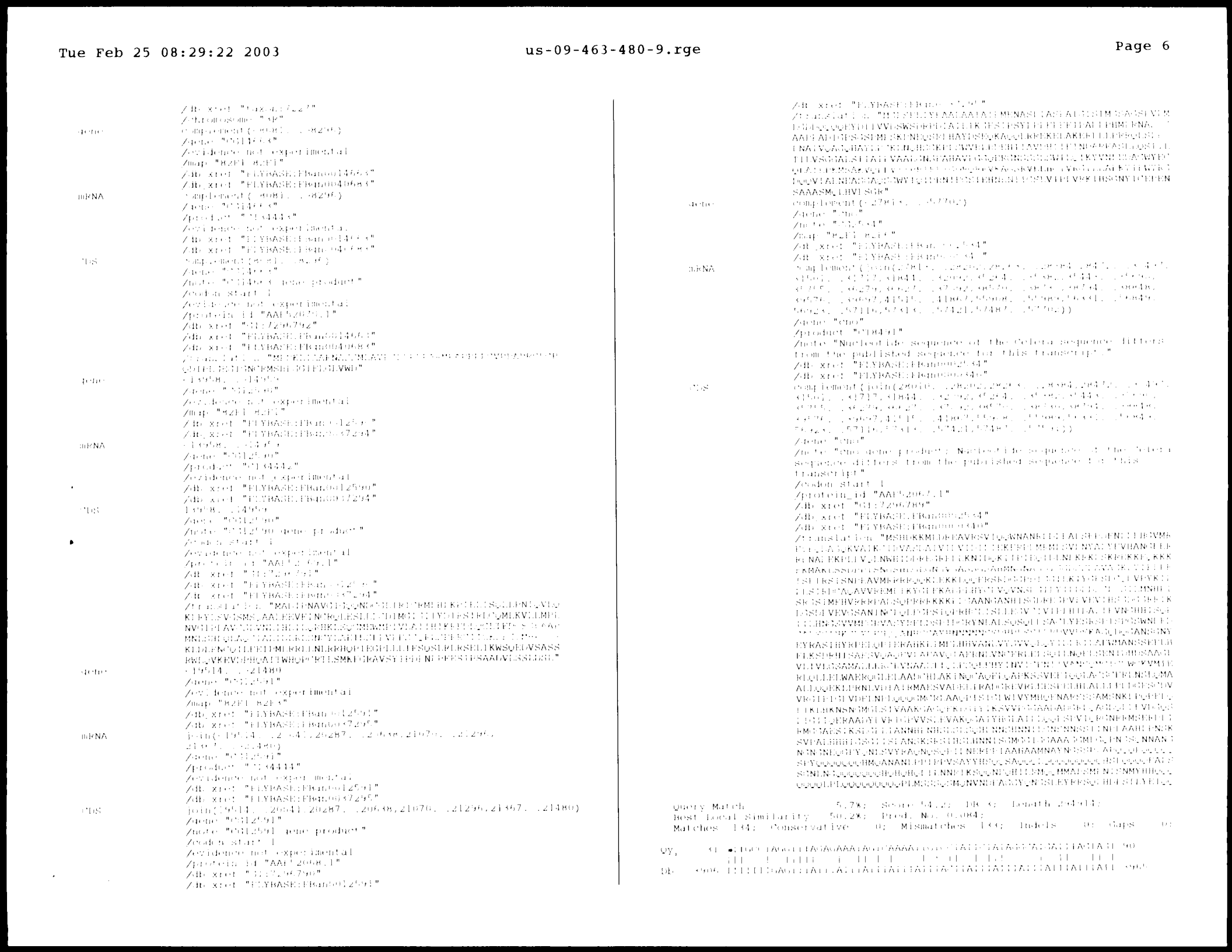
RESULT 1
AF110779
LOCUS AF110779 625 bp mRNA linear PLN 12-MAR-1999
DEFINITION Lilium longiflorum LSC1 (LSC1) mRNA, complete cds.
ACCESSION AF110779
VERSION AF110779.1 GI:4406381
KEYWORDS .
SOURCE Lilium longiflorum.
ORGANISM Lilium longiflorum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
Lilium.
REFERENCE 1 (bases 1 to 625)
AUTHORS Xu,B., Svoboda,I., Bhalla,P.L. and Singh,M.B.
TITLE Male gametic cell specific gene expression in flowering plants

```
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 96 (5), 2554-2558 (1999)
MEDLINE    [illegible]
PUBMED     [illegible]
REFERENCE  2  (bases 1 to [illegible])
  AUTHORS  Xu,H., Swoboda,I., Bhalla,P.L. and Singh,M.B.
  TITLE    Direct Submission
  JOURNAL  Submitted (03-DEC-1998) Environmental Horticulture and Resource
           Management, University of Melbourne, Parkville, Victoria 3052,
           Australia
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Lilium longiflorum"
                     /db_xref="taxon:[illegible]"
                     /sex="male"
                     /cell_type="gametes"
     gene            1..[illegible]
                     /gene="LGC1"
     CDS             82..468
                     /gene="LGC1"
                     /note="male gametic cell specific"
                     /codon_start=1
                     /product="LGC1"
                     /protein_id="[illegible]"
                     /db_xref="[illegible]"
                     /translation="MAAVAV[illegible]MVEKAALAL[illegible]
                     [illegible]
                     AILVA[illegible]KTAW[illegible]"
BASE COUNT      218 a    11[illegible]    1[illegible]    181 t
ORIGIN

  Query Match              [illegible]%;  Score [illegible];  DB [illegible];  Length 625;
  Best Local Similarity    [illegible]%;  Pred. No. 2.[illegible];
  Matches  11[illegible];  Conservative    [illegible];  Mismatches   14;  Indels    [illegible];  Gaps    2;

Qy        817 [illegible] 876
              ||||| |.||||||.||| |||.||||||||| ||||| || |||| |||||| ||||||
Db          2 CCATCCCATCAACAGAAGTTTAAGTGGAAATCCATTTCATTAGAAAAGATGGACAAAG 61

Qy        877 [illegible]
              ||||  [illegible]  ||||||||||||||||||
Db         62 [illegible] 121

Qy        934 GTATGGTTCACA 945
              |.|||  |  |||
Db        122 GTATGGTTCACA 133


RESULT 2
AC115575/c
LOCUS       AC115575               27291 bp    DNA     linear   HTG 21-MAR-2002
DEFINITION  Dictyostelium discoideum chromosome 2 map 1180800-1208089 strain
            AX4, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION   AC115575
VERSION     AC115575.1  GI:[illegible]
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 27291)
  AUTHORS   Gloeckner,G., [illegible], Szafranski,K., [illegible], [illegible],
            Lehmann,R., Baumgart,C., [illegible], April,J.F., [illegible], Kampf,K.,
            [illegible], Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
            Noegel,A.A.
  TITLE     Sequence and Analysis of Chromosome 2 of Dictyostelium
  JOURNAL   Unpublished
  REMARK    The Dictyostelium Genome Sequencing Consortium
REFERENCE   2  (bases 1 to 27291)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT     CDS predictions from GeneID may contain errors. Further Information
            is available from IMB Jena, Department of Genome Analysis
            (http://genome.imb-jena.de/dictyostelium/)
            and the University Cologne, Institute for Biochemistry I
            (http://www.uni-koeln.de/[illegible]/project.shtml
            Funding
            Agency : Deutsche Forschungsgemeinschaft (DFG).
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1..27291
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /chromosome="2"
                     /map="1180800-1208089"
     CDS             complement(join(1[illegible]..[illegible],12[illegible]..[illegible]))
                     /note="ORF_ID:dd_00011"
                     /codon_start=1
                     /product="Hypothetical protein ydeW precursor"
                     /protein_id="AAL92187.1"
                     /db_xref="GI:19569868"
                     /translation="MKY[illegible]
                     [illegible]
                     [illegible]
                     [illegible]
                     [illegible]NEMI[illegible]
                     INY[illegible]PQYLAP[illegible]
                     MY[illegible]
                     KIWMHLGY[illegible]
                     N"
     CDS             complement(join(2838..[illegible],3546..[illegible]))
                     /note="ORF_ID:dd_00013"
                     /codon_start=1
                     /product="[illegible]"
                     /protein_id="AAL92188.1"
                     /db_xref="GI:19569869"
                     /translation="M[illegible]
                     [illegible]
                     VWEDNGEHVLNF[illegible]NNNNNNN
                     NNNNNNNNNNN[illegible]LA"
     CDS             join(3837..3886,4358..[illegible])
                     /note="ORF_ID:dd_00015"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL92189.1"
                     /db_xref="GI:19569870"
                     /translation="M[illegible]KINNKL
                     [illegible]
                     [illegible]
                     [illegible]
                     [illegible]
                     NNGGKLEENCDANFLIDIY"
     CDS             join(6320..6707,6802..7274)
                     /note="ORF_ID:dd_00017"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL92190.1"
                     /db_xref="GI:[illegible]"
                     /translation="M[illegible]
                     [illegible]
                     [illegible]
                     [illegible]
                     [illegible]
                     CDPIEVCKIH"
     CDS             join(7879..8275,8373..8845)
                     /note="ORF_ID:dd_00018"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL92191.1"
                     /db_xref="GI:[illegible]"
                     /translation="M[illegible]
                     [illegible]NNA
```

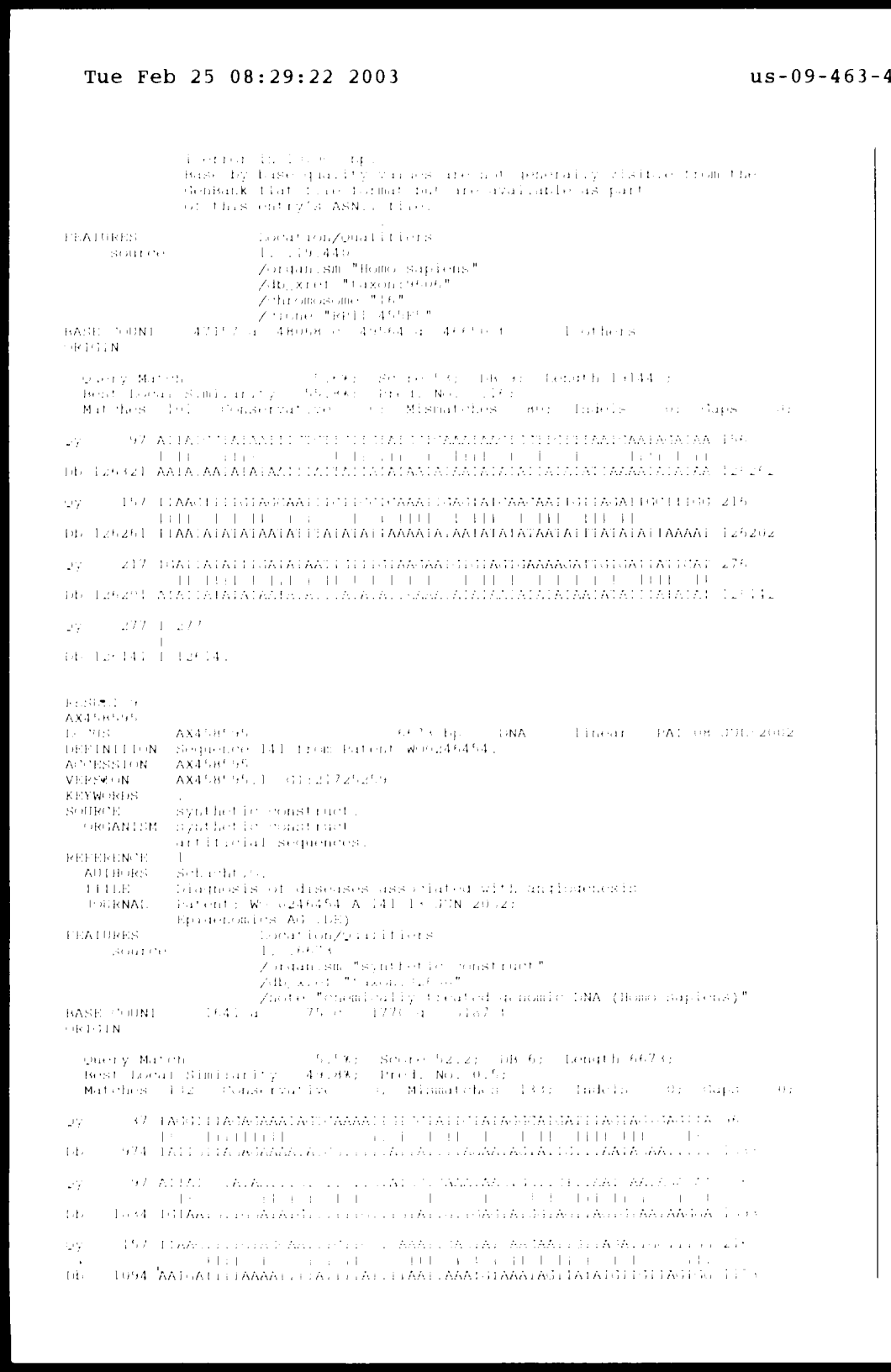
```
JOURNAL     Unpublished
REMARK      The Dictyostelium Genome Sequencing Consortium
REFERENCE   2  (bases 1 to 25117)
AUTHORS     Baumgart,C.
TITLE       Direct Submission
JOURNAL     Submitted (25-MAR-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT     CDS predictions from GeneID may contain errors. Further Information
            is available from IMB Jena, Department of Genome Analysis
            [illegible]
            and the University Cologne, Institute for Biochemistry I
            [illegible]
            Funding
            Agency : Deutsche Forschungsgemeinschaft (DFG).
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1..25117
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /chromosome="2"
                     /map="2753614-2779768"
     CDS             1118..3438
                     /note="ORF_ID:dd_01630"
                     /pseudo
                     /codon_start=1
     CDS             join(4048..4501,4579..6506)
                     /note="ORF_ID:dd_01629"
                     /codon_start=1
                     /product="Putative Chromosome Segregation ATPase"
                     /protein_id="AAL93602.1"
                     /db_xref="GI:19703140"
                     /translation="[illegible]
                     NSLSVCAIVIVIACAIIAIIMSIVRYEERAFFLGMLAKAKLSKISKKLSKK"
     CDS             join(6826..7002,7125..7262,7299..7632,7712..8487)
                     /note="ORF_ID:dd_01628"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL93603.1"
                     /db_xref="GI:19703141"
                     /translation="[illegible]
                     FIIWGSNYIDAISTFIGKIISSCIEVSDYKKCKKIGPEELSHPFKLIPSQSDHSIFF
                     KIEK"
     CDS             join(9165..9399,9463..9764)
                     /note="ORF_ID:dd_01627"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL93604.1"
                     /db_xref="GI:19703142"
                     /translation="[illegible]
                     PSIKIHGF"
     CDS             join(11059..11343,11444..11516,11628..11814,12083..
                     12768..14864)
                     /note="ORF_ID:dd_01626"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL93605.1"
                     /db_xref="GI:19703143"
                     /translation="[illegible]"
     CDS             complement(join(15086..15705,15982..16546))
                     /note="ORF_ID:dd_01625"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL93606.1"
                     /db_xref="GI:19703144"
                     /translation="[illegible]
                     NK"
     CDS             complement(join(17079..19515,19709..20690))
                     /note="ORF_ID:dd_02339"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL93607.1"
                     /db_xref="GI:19703145"
                     /translation="[illegible]"
     CDS             complement(join(22671..22910,23013..23359,23480..23555))
                     /note="ORF_ID:dd_02340"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL93608.1"
                     /db_xref="GI:19703146"
                     /translation="[illegible]
```

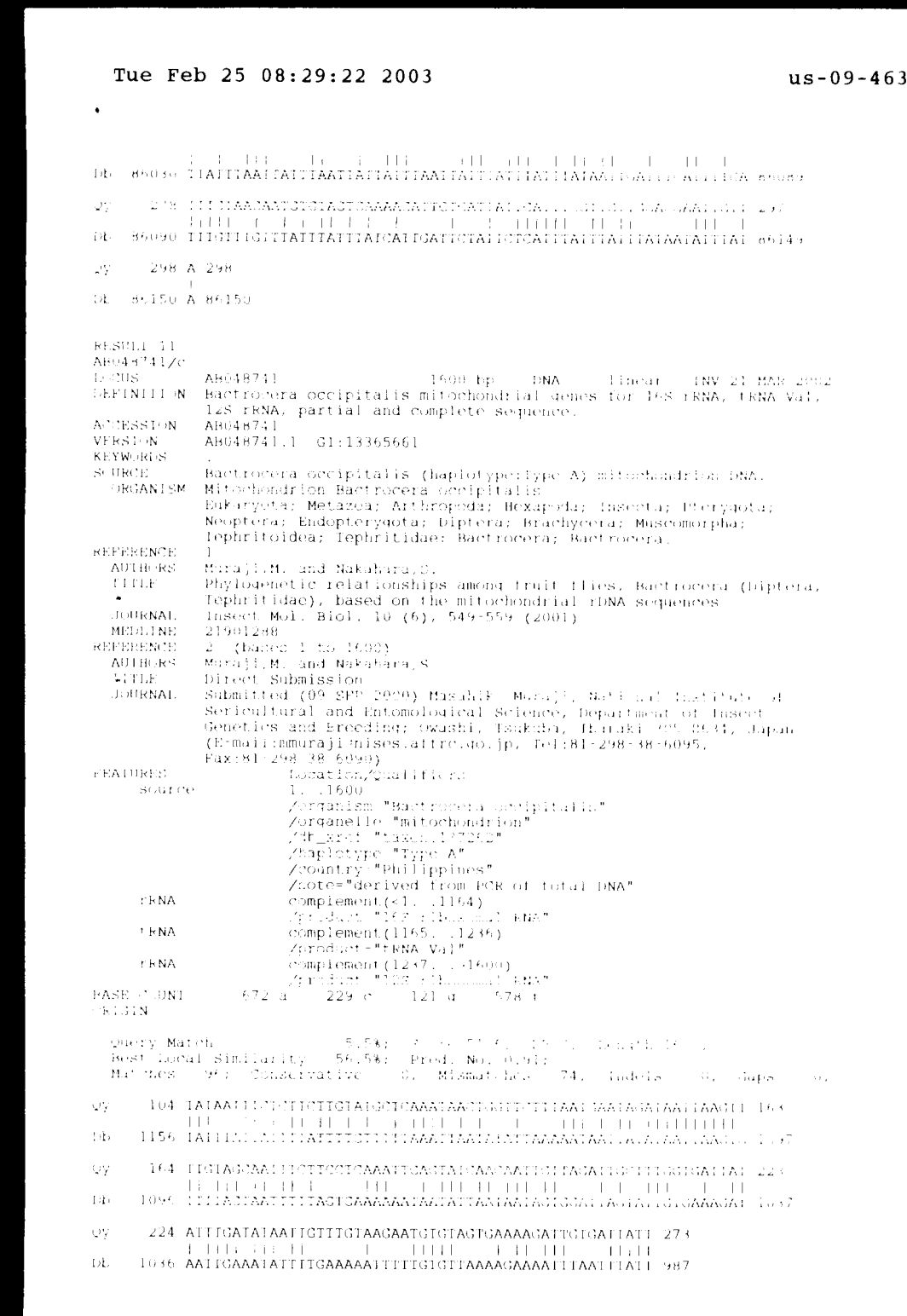
```
                     YETGFVESVLSIRGLLLSLSNIGLVLIESALSLYRAALLIRALIALSLLSAGIM
                     GNAAFKEISFANQTFLFNFSTIGAVFLLIAIIVVLGLPTAFLLSAFGLGLLG
                     SQVDPMCGGDIYKKFIGSHTIDEDCSKLSWGPTTAWIAIVVSLGIAFPNNTFVYCGPVF
                     KS"
     CDS             24236..25094
                     /note="ORF_ID:dd_02441"
                     /pseudo
                     /codon_start=1
BASE COUNT    10457 a   2498 c   2647 g   9525 t
ORIGIN

  Query Match           5.9%;  Score 55.8;  DB 2;  Length 25117;
  Best Local Similarity   52.3%;  Pred. No. 0.068;
  Matches  123;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy      44 AGAGAAATAGTCAAAATTCTCCTATTCTATAGCCATGATTTAGTAGTGAGTTAATTAATCC 103
           | | ||| |  ||||   |   ||| |||       ||||  || | | ||| ||||
Db   19676 AAAAAAAAAAACAAAAAATTTAATTTTATTAATTATATTTTTATTTATTTATTTATTT 19617

Qy     104 TATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAGTT 163
           || | || | |  || | |  | |  || |  || |||| | | |   | ||| | ||
Db   19616 AATTAATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 19557

Qy     164 TTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTAT 223
           | ||   |||| ||  |  | ||     |  || || | ||| ||       | |    ||
Db   19556 ATTTATTTATTTATTAATTTATTTATTAAAAAAAAAGTATTAAATGAAGAAGATTCAAAT 19497

Qy     224 ATTTGATATAATTGTTTGTAACAATGTGTAGTGAAAAGATTGTCATTATTCATTT 278
           |||  |  ||| |  || | |||| ||| | |    | |     ||| || || |
Db   19496 ATTAAAGTTAAATCATTATTAGAAAGTGGATTAGGTAAAGGTCAATTTTTAATAT 19442


RESULT 4
AX251284
LOCUS       AX251284               7814 bp    DNA     linear   PAT 05-OCT-2001
DEFINITION  Sequence 252 from Patent WO0168912.
ACCESSION   AX251284
VERSION     AX251284.1  GI:15984707
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 7814)
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with tumor suppressor genes and
            oncogenes
  JOURNAL   Patent: WO 0168912-A 252 20-SEP-2001;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1..7814
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT     1677 a    101 c   1779 g   4257 t
ORIGIN

  Query Match           5.8%;  Score 54.6;  DB 6;  Length 7814;
  Best Local Similarity   49.5%;  Pred. No. 0.15;
  Matches  141;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy      13 AATTAGGTTTGCCTAGGGTTTGCCTAGGTTTAGAGAAATAGTCAAAATTGTCCTATTCTA 72
           | ||,  | |   || | ||          || |   | | || |  | | |  || | ||
Db    3712 ATTTAATTATTTATATGTTTATTAATTAGTTCGTTTATTTGTTATTAGTTTATTAGTTTA 3771

Qy      73 TAGGCATGATTTAGTAGTGAGTTAATTATCCTATAATTTCTCTTCTTGTTATGCCTCAAATA 132
           ||    |  ||   || | | |||   ||    | |||| | |  || | || | |  ||
Db    3772 TAATTTTATTTGTTTTATTTATTTATGAATTTGTTTATTTATTTATTTATTTGTTTATTTA 3831

Qy     133 ACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTTCTTCCTCAAATTGAGTA 192
           | ||  ||| | | |  | ||| | ||  ||   |||| ||  | || ||   ||
Db    3832 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTGTTTATTAAATTTATTTA 3891

Qy     193 TTAACAATTGTTAGATTGCTTTGGTGATTATATTTGATATAATTGTTTGTAACAATGTGT 252
```

```
           | |  ||| |  || ||| | |||  ||  ||| | | | | | |  || | |
Db    3892 TTATTAATTTATTTATTTTTATTTATTTATTTATTTATTTATTTATTTATTTATT 3951

Qy     253 AGTGAAAAGATTGTCATTATTCATTTCGTTGTTGGACGAATTGTT 297
           ||    | |  ||||| ||| ,| ||     | || ||
Db    3952 TTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATT 3996


RESULT 5
AE003603
LOCUS       AE003603               294914 bp    DNA     linear   INV 04-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386043 section 5
            of 8, complete sequence.
ACCESSION   AE003603
VERSION     AE003603.2  GI:10727164
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 294914)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  MEDLINE   20196006
   PUBMED   10731132
REFERENCE   2  (bases 1 to 294914)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     On Oct 9, 2000 this sequence version replaced gi:7296751.
FEATURES             Location/Qualifiers
     source          1..294914
                     /organism="Drosophila melanogaster"
```



```
Qy      91 GAGTTAATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAA 150
           | ||| ||||    | |||| | |  || | |  | |  || |  || |||| | | |
Db    3966 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATT 4025

Qy     151 AGATAAATTAATTTTTGTATTAA-TTTTCTC-AAATTGAGTATCAACAATTGTTAGATTGCTTTGG 210
           | ||| | || | ||  |||| ||  | |   | | ||| |  | || ||   ||
Db    4026 TATTTATTTATTTATTTATTTATTTATTTATTATTATTATTATTATTTATTTATTTA 4085

Qy     211 CTTTGGTGATTATATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCAT 270
           || | |  |||| | |   || | | | | ||   || | |      |    | |  ||
Db    4086 TTTACTTATTTATTTATTTAATTAATTTATTTATTTATTTATTTATTTATTTATTT 4145

Qy     271 ATTCATTTCTTGTTGGACGAATTGTT 297
           | |||'     ||     | || ||
Db    4146 ATTTATTTATTAATTTATTTATTTATT 4172


RESULT 5
AX346662/c
LOCUS       AX346662                7851 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 1733 from Patent WO0200928.
ACCESSION   AX346662
VERSION     AX346662.1  GI:18494548
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 1733 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1..7851
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT     1799 a     284 c    2146 g    3622 t
ORIGIN

  Query Match             5.6%;  Score 56;  DB 6;  Length 7851;
  Best Local Similarity   55.8%;  Pred. No. 0.33;
  Matches  101;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy      97 ATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAA 156
           | ||'  ||||    | | || || |||  |  ||||  |  |    |      |||| ||||
Db     911 AATATAATATATAGATTTATTATTATATAATATAAATATATTATATATTAAAATATATAA 855

Qy     157 TTAAGTTTTGTATCAATTTCTTGCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGG 216
           ||||  | | ||  | ||   | | |||| | |||  | |||  ||| ||
Db     854 TTAATATATATAATATTTATATATTAAAATATAAATATATATAATATTTATATATTAAAAT 795

Qy     217 TGATTATATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCAT 276
           ||||||| | ;|||| || | | | | | ||| ' || |  | | | | |  |||  ||
Db     794 ATATTATATATAATATATTTATATATTAAAATATATATATAATATATTTATATATTAT 735

Qy     277 T 277
           |
Db     734 T 734


RESULT 7
AX346663
LOCUS       AX346663                7851 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 1734 from Patent WO0200928.
ACCESSION   AX346663
VERSION     AX346663.1  GI:18494549
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 1734 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1..7851
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT     1811 a     284 c    2095 g    3661 t
ORIGIN

  Query Match             5.6%;  Score 56;  DB 6;  Length 7851;
  Best Local Similarity   55.8%;  Pred. No. 0.33;
  Matches  101;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy      97 ATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAA 156
           | ||| ||||    | | || || |||  |  ||||  |  |    |      |||| ||||
Db    6938 AATATAATATATAGATTTATTATTATATAATATAAATATATTATATATTAAAATATATAA 6997

Qy     157 TTAAGTTTTGTATCAATTTCTTGCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGG 216
           ||||  | | ||  | ||   | | |||| | |||  | |||  ||| ||
Db    6998 TTAATATATATAATATTTATATATTAAAATATAAATATATATAATATTTATATATTAAAAT 7057

Qy     217 TGATTATATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCAT 276
           ||||||| | ;|||| || | | | | | ||| ' || |  | | | | |  |||  ||
Db    7058 ATATTATATATAATATATTTATATATTAAAATATATATATAATATATTTATATATTAT 7117

Qy     277 T 277
           |
Db    7118 T 7118


RESULT 8
AC012645/c
LOCUS       AC012645              191440 bp    DNA     linear   PRI 06-MAR-2002
DEFINITION  Homo sapiens chromosome 16 clone RP11-455E5, complete sequence.
ACCESSION   AC012645
VERSION     AC012645.6  GI:19172854
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 191440)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191440)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 191440)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Mar 6, 2002 this sequence version replaced gi:9256433.
            Sequence Quality Assessment:
            This entry has been annotated with sequence quality
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than
            1 error in 10,000 bp.
            Base-by-base quality values are not generally visible from the
            GenBank flat file format but are available as part
            of this entry's ASN.1 file.
            ---- . ----------
            Sequence Quality Assessment:
            This entry has been annotated with sequence quality
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than
```





```
RESULT 11
AB048741/c
LOCUS       AB048741                1600 bp    DNA     linear   INV 21-MAR-2002
DEFINITION  Bactrocera occipitalis mitochondrial genes for 16S rRNA, tRNA Val,
            12S rRNA, partial and complete sequence.
ACCESSION   AB048741
VERSION     AB048741.1  GI:13365661
KEYWORDS    .
SOURCE      Bactrocera occipitalis (haplotype:type A) mitochondrial DNA.
  ORGANISM  Mitochondrion Bactrocera occipitalis
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Tephritoidea; Tephritidae; Bactrocera; Bactrocera.
REFERENCE   1
  AUTHORS   Muraji,M. and Nakahara,S.
  TITLE     Phylogenetic relationships among fruit flies, Bactrocera (Diptera,
            Tephritidae), based on the mitochondrial rDNA sequences
  JOURNAL   Insect Mol. Biol. 10 (6), 549-559 (2001)
  MEDLINE   21901288
REFERENCE   2  (bases 1 to 1600)
  AUTHORS   Muraji,M. and Nakahara,S
  TITLE     Direct Submission
  JOURNAL   Submitted (09 SEP 2000) Masahiko Muraji, National Institute of
            Sericultural and Entomological Science, Department of Insect
            Genetics and Breeding; Owashi, Tsukuba, Ibaraki 305-8634, Japan
            (E-mail:mamuraji@nises.affrc.go.jp, Tel:81-298-38-6095,
            Fax:81-298-38-6099)
FEATURES             Location/Qualifiers
     source          1..1600
                     /organism="Bactrocera occipitalis"
                     /organelle="mitochondrion"
                     /db_xref="taxon:127262"
                     /haplotype="Type A"
                     /country="Philippines"
                     /note="derived from PCR of total DNA"
     rRNA            complement(<1..1164)
                     /product="16S ribosomal RNA"
     tRNA            complement(1165..1236)
                     /product="tRNA Val"
     rRNA            complement(1237..>1600)
                     /product="12S ribosomal RNA"
BASE COUNT      672 a    229 c    121 g    578 t
ORIGIN

  Query Match             5.5%;  Score 51.6;  DB 3;  Length 1600;
  Best Local Similarity   56.5%;  Prob. No. 0.91;
  Matches   96;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy      104 TATAATTTCTCTTTGTTGTAGCTCAAATAACTAATTTTTTAATTAATAGATAATTAAGTTT 163
            ||| |  | || || | | | |  |   |  || | || ||||||||
Db     1156 TATTTCTAATTTATTTTTTTTAAATTAATTATTAAAAATAATTAATAGATAATTAAGTTT 1097

Qy      164 TTGTAGGAATTTCTTCCTCAAATTCATTATCAATAATTCTTAGATTGTTTTGTTGATTAT 223
            || ||| | || |    |||   | ||| || ||| ||    | |  |||   |  ||
Db     1096 TTTTATTAATTTTTAGTCAAAAAATAATATTAATAATTATTGAATTATTATTGTAAAGAT 1037

Qy      224 ATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAACATTGTGATTATT 273
            | ||| || ||      |    |||||   | || |||    ||||
Db     1036 AATTGAAATATTTTGAAAAATTTTTGTGTTAAAAGAAAATTTAATTTATT 987
```

```
RESULT 12
AB048742/c
LOCUS       AB048742                1600 bp    DNA     linear   INV 21-MAR-2002
DEFINITION  Bactrocera occipitalis mitochondrial genes for 16S rRNA, tRNA Val,
            12S rRNA, partial and complete sequence.
ACCESSION   AB048742
VERSION     AB048742.1  GI:13365662
KEYWORDS    .
SOURCE      Bactrocera occipitalis (haplotype:type B) mitochondrial DNA.
  ORGANISM  Mitochondrion Bactrocera occipitalis
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Tephritoidea; Tephritidae; Bactrocera; Bactrocera.
REFERENCE   1
  AUTHORS   Muraji,M. and Nakahara,S.
  TITLE     Phylogenetic relationships among fruit flies, Bactrocera (Diptera,
            Tephritidae), based on the mitochondrial rDNA sequences
  JOURNAL   Insect Mol. Biol. 10 (6), 549-559 (2001)
  MEDLINE   21901288
REFERENCE   2  (bases 1 to 1600)
  AUTHORS   Muraji,M. and Nakahara,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09 SEP 2000) Masahiko Muraji, National Institute of
            Sericultural and Entomological Science, Department of Insect
            Genetics and Breeding; Owashi, Tsukuba, Ibaraki 305-8634, Japan
            (E-mail:mamuraji@nises.affrc.go.jp, Tel:81-298-38-6095,
            Fax:81-298-38-6099)
FEATURES             Location/Qualifiers
     source          1..1600
                     /organism="Bactrocera occipitalis"
                     /organelle="mitochondrion"
                     /db_xref="taxon:127262"
                     /haplotype="Type B"
                     /country="Philippines"
                     /note="derived from PCR of total DNA"
     rRNA            complement(<1..1164)
                     /product="16S ribosomal RNA"
     tRNA            complement(1165..1236)
                     /product="tRNA Val"
     rRNA            complement(1237..>1600)
                     /product="12S ribosomal RNA"
BASE COUNT      672 a    227 c    121 g    580 t
ORIGIN

  Query Match             5.5%;  Score 51.6;  DB 3;  Length 1600;
  Best Local Similarity   56.5%;  Prob. No. 0.91;
  Matches   96;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy      104 TATAATTTCTCTTTGTTGTAGCTCAAATAACTAATTTTTTAATTAATAGATAATTAAGTTT 163
            ||| |  | || || | | | |  |   |  || | || ||||||||
Db     1156 TATTTCTAATTTATTTTTTTTAAATTAATTATTAAAAATAATTAATAGATAATTAAGTTT 1097

Qy      164 TTGTAGGAATTTCTTCCTCAAATTCATTATCAATAATTCTTAGATTGTTTTGTTGATTAT 223
            || ||| | || |    |||   | ||| || ||| ||    | |  |||   |  ||
Db     1096 TTTTATTAATTTTTAGTCAAAAAATAATATTAATAATTATTGAATTATTATTGTAAAGAT 1037

Qy      224 ATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAACATTGTGATTATT 273
            | ||| || ||      |    |||||   | || |||    ||||
Db     1036 AATTGAAATATTTTGAAAAATTTTTGTGTTAAAAGAAAATTTAATTTATT 987
```

```
RESULT 13
AB048752/c
LOCUS       AB048752                1602 bp    DNA     linear   INV 21-MAR-2002
DEFINITION  Bactrocera correcta mitochondrial genes for 16S rRNA, tRNA Val, 12S
            rRNA, partial and complete sequence.
ACCESSION   AB048752
VERSION     AB048752.1  GI:13365672
KEYWORDS    .
SOURCE      Bactrocera correcta (haplotype:type A) mitochondrial DNA.
  ORGANISM  Mitochondrion Bactrocera correcta
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
```

```
             Tephritoidea; Tephritidae; Bactrocera; Bactrocera.
REFERENCE    1
  AUTHORS    Muraji,M. and Nakahara,S.
  TITLE      Phylogenetic relationships among fruit flies, Bactrocera (Diptera,
             Tephritidae), based on the mitochondrial rDNA sequences
  JOURNAL    Insect Mol. Biol. 10 (6), 549-559 (2001)
  MEDLINE    21901288
REFERENCE    2  (bases 1 to 1602)
  AUTHORS    Muraji,M. and Iwahashi,O.
  TITLE      Direct Submission
  JOURNAL    Submitted (05-SEP-2000) Masahiko Muraji, National Institute of
             Sericultural and Entomological Science, Department of Insect
             Genetics and Breeding; Owashi, Tsukuba, Ibaraki 305-8634, Japan
             (E-mail:mmuraji@nises.affrc.go.jp, Tel:81-298-38-6095,
             Fax:81-298-38-6099)
FEATURES             Location/Qualifiers
     source          1..1602
                     /organism="Bactrocera correcta"
                     /organelle="mitochondrion"
                     /db_xref="taxon:47773"
                     /haplotype="Type A"
                     /country="Thailand, Bangkok"
                     /note="derived from PCR of total DNA"
     rRNA            complement(<1..1166)
                     /product="16S ribosomal RNA"
     tRNA            complement(1167..1238)
                     /product="tRNA-Val"
     rRNA            complement(1239..>1602)
                     /product="12S ribosomal RNA"
BASE COUNT      672 a    225 c    123 g    582 t
ORIGIN

  Query Match             5.5%;  Score 51.6;  DB 3;  Length 1602;
  Best Local Similarity   56.5%;  Pred. No. 0.91;
  Matches   96;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy        104 TATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAGTT 163
              |||    ||   || || | |  | |||| | ||   |     ||| | || |||||||||
Db       1158 TATTTATTAATTTATTTTGTTTTTAAATTAATGTATTAAAAATAATTATATAATTAAGTG 1099

Qy        164 TTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTAT 223
              || |||    ||| |       |||    | | | || ||| ||||   | |  |||    | ||
Db       1098 TTTTAGTAGTTTTTTAATGAAAAAAATAATTTTAATAATAGTTAATTAGTATTGTGAAAGAT 1039

Qy        224 ATTTCATATAATTCTTCTTAATAATCTGTATTCAAAATATTTTGATTATT 273
              | || | ||| ||        |     ||| |      | || |||   | |||||
Db       1038 AATTGAAATATTTTTGAAAAATTTTTGTATTAAAAGAAAATTTAGTTTATT 989


RESULT 14
AF465307
LOCUS       AF465307                6281 bp    DNA     linear   INV 05-MAR-2002
DEFINITION  Dictyostelium discoideum ABC transporter ABCA.5 (abca5) gene,
            complete cds.
ACCESSION   AF465307
VERSION     AF465307.1  GI:19110823
KEYWORDS    .
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 6281)
  AUTHORS   Anjard,C. and Loomis,W.F.
  TITLE     Evolution of the ABC transporters of Dictyostelium
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 6281)
  AUTHORS   Anjard,C. and Loomis,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2002) Biology Dept. 0368, UCSD, 9500 Gilman
            Drive, La Jolla, CA 92093, USA
FEATURES             Location/Qualifiers
     source          1..6281
                     /organism="Dictyostelium discoideum"
                     /strain="Ax4"
                     /db_xref="taxon:44689"
     gene            <400..>5537
                     /gene="abca5"
     mRNA            join(<400..532,613..5077,5147..>5537)
                     /gene="abca5"
                     /product="ABC transporter ABCA.5"
     CDS             join(400..532,613..5077,5147..5537)
                     /gene="abca5"
                     /note="family A"
                     /codon_start=1
                     /product="ABC transporter ABCA.5"
                     /protein_id="AAL85298.1"
                     /db_xref="GI:19110824"
                     /translation="M[illegible]IVF[illegible]VF[illegible]VFAIL
                     VLVQLFRPN[illegible]LIR[illegible]NE[illegible]Y[illegible]NK
                     S[illegible]LVFAVTFKL[illegible]KGLMLNTFQINRTQVF[illegible]VANI[illegible]NSFKY
                     STRLDSNEVLF[illegible]NEVF[illegible]FNWAVL[illegible]PY
                     PDPYVELWQKW[illegible]NGELSVIKSAGSVFVSAALLIFIFERLVTEIVVFFFTKTF[illegible]MG
                     LNQYCYFISWITTSLVTALPIDLTTIVILKGSQVIHSTSWITVIVTTTLYL[illegible]QLLA
                     [illegible]PSMFFDFSKFA[illegible]FN[illegible]IN[illegible]GFYFTINTHTKLL[illegible]HSF[illegible]IACSF
                     YIMG[illegible]TPDT[illegible]TNLNQVI[illegible]VINT[illegible]WY[illegible]FEPW
                     YTLPTTS[illegible]W[illegible]NEENENE[illegible]YNYN[illegible]NNNN[illegible]NTFMV[illegible]VTNKTTISTK
                     NLRKFFK[illegible]AVN[illegible]MPF[illegible]KSTTTSMLTGLTSPISGIAF
                     TLGN[illegible]TTQMNE[illegible]IWN[illegible]IVL[illegible]YA[illegible]NKK[illegible]KM
                     AN[illegible]VWDF
                     LLCHRK[illegible]TY[illegible]ALTL[illegible]FNKY[illegible]SK
                     S[illegible]PNV[illegible]VT[illegible]PDVTVLSNA[illegible]IVE[illegible]KDFDENFS
                     NFHIDSY[illegible]ISVTTLEEVFLKIGT[illegible]DAL[illegible]VSS[illegible]FQQ
                     LKALLIK[illegible]VIVT[illegible]VT[illegible]NEVTFYNNSIEPLTFSL
                     TQQYGKNDIVF[illegible]QLAGTTSEFDFNKTLSNGFYTNQF[illegible]YLNN[illegible]NYKQSSS
                     STNFTTPLSSSLINNTTKVTSYNSL[illegible]NY[illegible]VHVNLTNEALL[illegible]NHNGIGIEC[illegible]S
                     LPFDHITLISFQKASQGMNI[illegible]IVYFTVIMMAGFSLMA[illegible]NRIKRLTYT
                     SGCKKYTYWVSNLLWDYFFAFILLLFICTTLAIVDDKFNDQFGLFFLSLVLFSLSTTF
                     LSYLFS[illegible]GA[illegible]AT[illegible]GVIMT[illegible]AMT[illegible]TEVT[illegible]KN[illegible]IAFTIDT
                     VFNILSPLYAFSRVIFTTSGFPGSLRLGALKTDDYWSLHYASTPTTTLSVHVVVWLFF
                     TLLIDYSPEFKGYLRNFKTTLFFFPFIDELSDVSAERTRLESMSPTPSTDQGSGGGD
                     MLQYK[illegible]HELP[illegible]KNENAVYNST[illegible]VCMLAGDIL
                     PISGQITINGHDLVTDRAQALRGVSVCPQFDALTTLLTAREQLSLYCAIKGVPDDKIS
                     LVVEAFIEMM[illegible]KIANSNT[illegible]NFFKVS[illegible]IAM[illegible]NFSVVLIDEASSGCDPTTR
                     YKQ[illegible]QVISHL[illegible]KNKVTILTSHSMCE[illegible]QAL[illegible]RMTIMM[illegible]FKC[illegible]IKG[illegible]KIGAGY
                     SVEVKFKKS[illegible]LEV[illegible]SL[illegible]VL[illegible]FNA[illegible]DQHDLMA[illegible]NE[illegible]KVSETEN
                     TLSTELSSTLDDYSVSQTSLEQVFLKLTGATHEDRLNLNNQQQQQQTTPNSD"
BASE COUNT     2307 a    726 c    857 g   2391 t
ORIGIN

  Query Match             5.4%;  Score 51.4;  DB 3;  Length 6281;
  Best Local Similarity   51.1%;  Pred. No. 0.74;
  Matches  121;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy         42 TTAGAGAAATAGTCAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTAATTAT 101
              ||  |  |||  | | |||| |  |||     | |     |  ||    || | | | | ||||
Db        154 TTTTATTAATT[illegible]AATTTTAATTTAT 213

Qy        102 CCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAG 161
                    | | || | |  || | |  | |  ||     || |||| | | |   ||||| |
Db        214 TTATTTATTTATTTATTTATTTATTTATTAATTTATTTATTTATTTATTAATTTAT 273

Qy        162 TTTTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATT 221
              || | ||  |||| ||  |  | ||    |||  |     | |    |||    |   |  ||
Db        274 TTATTTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTAAATATTTTTT 333

Qy        222 ATATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCATTT 278
               | ||| ||| | || |    ||| ||       | | |||| | |    ||||    ||| |
Db        334 TTTTTTTATAAATTTTTGATTAAAAAAAAAAAATAAAAAAAATAATATTAAATATAT 390


RESULT 15
AC117074/c
LOCUS       AC117074               61052 bp    DNA     linear   HTG 06-APR-2002
DEFINITION *Dictyostelium discoideum chromosome 2 map 2779865-2840915 strain
            AX4, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION   AC117074
VERSION     AC117074.1  GI:20066241
```

```
KEYWORDS    HIG; HIGS_PHASE2.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE   1  (bases 1 to 61052)
  AUTHORS   Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Bean,P.,
            Lehmann,R., Baumgart,C., Parra,G., April,J.F., Cubo,R., Kumpf,K.,
            Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
            Noegel,A.A.
  TITLE     Sequence and Analysis of Chromosome 2 of Dictyostelium
  JOURNAL   Unpublished
  REMARK    The Dictyostelium Genome Sequencing Consortium
REFERENCE   2  (bases 1 to 61052)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-APR-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT     CDS predictions from GeneID may contain errors. Further Information
            is available from IMB Jena, Department of Genome Analysis
            (http://www.[illegible],project.shtml)
            and the University Cologne, Institute for Biochemistry I
            (http://www.[illegible],project.shtml
            Funding
            Agency : Deutsche Forschungsgemeinschaft (DFG).
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1..61052
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /chromosome="2"
                     /map="2779865 2840915"
BASE COUNT    25445 a   6805 c   6328 g  22474 t
ORIGIN

  Query Match          5.4%;  Score 51.4;  DB 2;  Length 61052;
  Best Local Similarity   51.1%;  Pred. No. 0.45;
  Matches  121;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy      42 TTAGAGAAATATTTAAAATTTCGTAATTATAGGGATGATTAGTAGTGATTAATTAT 101
           ||  |  |||  | | |||| |  |||| ||          ||   || | | ||| |||
Db    6576 TTTTATTAATTTTTATAATTTTTATATTGTAATATTTATTTTATTTATTTATTTATT 6517

Qy     102 CATATAATTTCTCTTCTTGTTATGTTCAAAATAATCGTTCTTTAATGAATAGATAATTAAG 161
           | |||| | |  || | |  | |  ||  |  || |||| | | |   | ||| |
Db    6516 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATT 6457

Qy     162 TTTGTTAGCAATTTCTTCCTCAAATTGAGTATTAACAATTGTTAGATTGCTTTGGTGATT 221
           || | ||   |||| ||  |  | ||   ||| |    | |   |||   |   | ||
Db    6456 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAATATTTTTTT 6397

Qy     222 ATATTTGATATAATTGTTTGTAAGAATGTGTATTGAAAAGATTGTGATTATTCATTT 278
           | ||| | | | || |   ||| ||     | | |||| | |   ||||   || |
Db    6396 TTTTTTATAAATTTTTGATTAAAAAAAAAAAAATAAAAAAAATAATTAAATATAT 6340


Search completed: February 24, 2003, 14:25:42
Job time : 2981 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February, 24, 2003, 10:08:36 ; Search time 272 Seconds
(without alignments)
7824.038 Million cell updates/sec

Title: US-09-463-480-9
Perfect score: 945
Sequence: 1 qaaaggtgttggaaattaaggt.....tctcttctgtatgattcaaca 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2195229 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters. 4370458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_181002.*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA2002.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqp-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	945	100.0	945	20	AAX19930	Lily LGC1 promoter
2	83.6	8.8	625	20	AAX19927	Lily LGC1 nucleoti
3	54.6	5.8	7814	22	AAS46530	Tumour suppressor
c 4	53	5.6	7851	24	ABL33760	Human immune syste
5	53	5.6	7851	24	ABL33761	Human immune syste
6	52.2	5.5	6673	24	ABQ67111	Human angiogenesis
7	51	5.4	6126	22	AAS46573	Tumour suppressor
8	51	5.4	6126	24	ABL33830	Human immune syste
9	51	5.4	8093	22	AAS46435	Tumour suppressor
10	51	5.4	8093	24	ABL92236	Chemically treated
11	51	5.4	8093	24	ABK22973	Human DNA for stud
12	51	5.4	8093	24	ABL49331	Human MLH1 gene as
13	50.6	5.4	15732	22	AAS45388	Chemically pretrea
14	50.6	5.4	15732	24	ABK82838	DNA transcription
15	50.4	5.3	11790	24	ABL32543	Human immune syste
16	50.2	5.3	5984	24	ABQ66994	Human angiogenesis
17	50.2	5.3	6861	24	ABL32741	Human immune syste
18	50.2	5.3	6668	22	AAS46118	Tumour suppressor
19	50.2	5.3	6668	24	ABN80162	Human chemically m
20	50.2	5.3	6668	24	ABL33219	Human immune syste
21	50	5.3	5070	24	ABL32801	Human immune syste
22	50	5.3	6286	22	AAS46591	Tumour suppressor
23	49.4	5.2	4045	24	ABD92420	Chemically treated
24	49.4	5.2	4045	24	AAD22341	Chemically treated
25	48.8	5.2	11022	24	ABL34105	Human immune syste
26	48.6	5.1	6904	24	ABL32212	Human immune syste
27	48.6	5.1	8446	24	ABL23671	Human immune syste
28	48.2	5.1	5535	24	ABL92216	Chemically treated
29	48.2	5.1	5535	24	ABL49317	Human polynucleoti
30	48.2	5.1	5535	24	ABL32664	Human immune syste
31	48.2	5.1	5987	24	ABQ67102	Human angiogenesis
32	48.2	5.1	5987	24	ABL33563	Human immune syste
33	47.8	5.1	6956	24	ABL70225	Chemically treated
34	47.8	5.1	7597	24	ABL33013	Human immune syste
35	47.8	5.1	8067	24	ABN80324	Human chemically m
36	47.8	5.1	8663	22	AAS46711	Tumour suppressor
37	47.6	5.0	6291	22	AAS46152	Tumour suppressor
38	47.6	5.0	7746	24	ABK40048	Human chemically p
39	47.6	5.0	7746	24	ABL33857	Human immune syste
40	47.4	5.0	11260	22	AAS45511	Chemically pretrea
41	47.4	5.0	11260	24	ABN80338	Human chemically m
42	47.4	5.0	11269	24	ABK82832	DNA transcription
43	47.4	5.0	12521	22	AAS46121	Tumour suppressor
44	47	5.0	6104	24	ABL33956	Human immune syste
45	47	5.0	6107	24	ABL33957	Human immune syste

ALIGNMENTS

RESULT 1
AAX19930
ID AAX19930 standard; DNA; 945 BP.
XX
AC AAX19930;
XX
DT 11-JUN-1999 (first entry)
XX
DE Lily LGC1 promoter nucleotide sequence.
XX
KW ...; Lilium longiflorum; LGC1; qHLA, qCH3, pollen, male; germ line;
KW promoter, sterile, transgenic, specific, pollination, seedless fruit;
KW pollen mother cell, gamete, histone; ss.
XX
OS Lilium longiflorum.
XX
PN WO9905281-A1.
XX
PD 04-FEB-1999.
XX
PF 24-JUL-1998; 98WO-AU00587.
XX
PR 31-DEC-1997; 97AU-0001184.
PR 25-JUL-1997; 97AU-0028233.
XX
PA (UYME) UNIV MELBOURNE.
XX
PI Bhalla P, Singh M, Swoboda I, Xu H;
XX
DR WPI, 1999-142936/12.
XX
PT New isolated nucleic acid for, e.g. generating male sterile plants -

```
PI   comprises a nucleotide sequence that allows expression of a
PI   transposase gene in generative and sperm cells of a plant
XX
PS   Claim 5; Fig 1; 64pp; English.
XX
CC   The present sequence represents the LAT1 promoter from lily (Lilium
CC   longiflorum). The sequence specifically allows expression of the
CC   gene in generative and sperm cells of a plant. Also described are: (1) a
CC   genetic construct comprising a generative cell and sperm cell specific
CC   promoter operably linked to a transposase, which is capable of inducing
CC   transposition of a transposable element such that upon expression of the
CC   promoter, the transposase gene is expressed facilitating transposition of
CC   the element; (2) inducing male sterility in a plant by operably linking a
CC   cytotoxic nucleic acid to a promoter, which directs male gamete specific
CC   expression to produce a product that the cytotoxic nucleic acid that
CC   inactivates or kills male gametes in the plant; and (3) a male sterile
CC   plant produced by the method. The male gamete promoter and/or gene is
CC   useful for facilitating male gamete specific transposon tagging, where
CC   the product of pollen grains is screened for a range of phenotypes of
CC   interest, and the transposon tagged plants are used to clone certain
CC   genes. The method is useful for producing seedless fruit or fruit with
CC   reduced seed content, particularly where pollination stimulates fruit
CC   development, and where the lack of fertilisation results in seedless
CC   fruit. The nucleic acids enable specific genetic manipulation of the
CC   male germline by generating male sterile plants, and facilitating male
CC   gamete specific transposon tagging.
XX
SQ   Sequence 945 BP; 269 A; 144 C; 220 G; 313 T; 0 other;

  Query Match             100.0%;  Score 945;  DB 20;  Length 945;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-222;
  Matches  945;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AAAGGAAATTTCGAATTAGTTTTTCTAAGTTTTCTTAGTGTTTTAGTAAAATAGTTAAAAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AAAGGAAATTTCGAATTAGTTTTTCTAAGTTTTCTTAGTGTTTTAGTAAAATAGTTAAAAA 60

Qy      61 TGTCCTACTTATAGCTATGCTATCTAGTGAGTTAATTATCTATAATTTCTCTTCTTG 120
           ||||||.||||.|||.|||||||||||||||||.||||||.|||||||||||||||||||
Db      61 TGTCCTATTCTATAGCTATGCTATCTAGTGAGTTAATTATCTATAATTTCTCTTCTTG 120

Qy     121 TATCTCAAATAATCTCTCTTTAAATAATATATAATAATTTTTTTTTTTTTTTTTTTTTTT 180
           |||||||.|||||.|||.|||||||.|.|||.|||||.|||.|||||||||||||||||||
Db     121 TATCTCAAATAATCTCTCTTTAAATAATATATAATAATTTTTTTTTTTTTTTTTTTTTTT 180

Qy     181 CAAATTGAGTATCAATAATTTTTATATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 240
           ||||||.||||.|||.|||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAAATTGAGTATCAATAATTTTTATATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 240

Qy     241 GTAAGAATGTGTAGTGAAAAGATTGGATTATTCATTGTTTGTTGGACGAATTGTTAGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTAAGAATGTGTAGTGAAAAGATTGGATTATTCATTGTTTGTTGGACGAATTGTTAGA 300

Qy     301 TTTCATCGTAATGTAATAGTATATGTTGAAAATATTTGTGGAATAGAAGATGAAAAATTT 360
           ||||||.|||.|.||.|||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTTCATCGTAATGTAATAGTATATGTTGAAAATATTTGTGGAATAGAAGATGAAAAATTT 360

Qy     361 TATTCTTTGTATTTGTTTAAAAAATTTTCTTTTCATTATTCCATTGAATGTTAACCACT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TATTCTTTGTATTTGTTTAAAAAATTTTCTTTTCATTATTCCATTGAATGTTAACCACT 420

Qy     421 TGTCATTCATCTGACTGGGATTCTCGACTCATCTAGTGATCGCTGCGATATTTGA 480
           |||||||.|||.|.|||.|.|||||.|||||||||||||||||||||||||||||||||
Db     421 TGTCATTCATCTGACTGGGATTCTCGACTCATCTAGTGATCGCTGCGATATTTGA 480

Qy     481 TTTTTTTTTTTAATAAATTTTCATAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 540
           |||||||||||.|||.|||||||.|||||||||||||||||||||||||||||||||||||
Db     481 TTTTGGTGTGT AATAAATTGTGATAATGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 540

Qy     541 TGATATGAATATGAGTTGATTATGTTAATGGATGGATATTTATCATGTGGTTGAATTT 600
           ||||||||||||.||||||||||||||.|||||||||||||||||||||||||||||||||
Db     541 TGATATGAATATGAGTTGATTATGTTAATGGATGGATATTTATCATGTGGTTGAATTT 600

Qy     601 GAGTGCTGTAATGAGATAGAATTTGAAAGTTCTAATAATAAACTTTAATTTGAATTTTGGC 660
           ||||||||.|||||||.|||||||||||||||||||||||||||||||||||||||||||||
Db     601 GAGTGCTGTAATGAGATAGAATTTGAAAGTTCTAATAATAAACTTTAATTTGAATTTTGGC 660

Qy     661 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 720

Qy     721 TATTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TATTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 780

Qy     781 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 840

Qy     841 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 900

Qy     901 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 945
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 945


RESULT 2
AAX19927
ID   AAX19927 standard; DNA; 945 BP.
XX
AC   AAX19927;
XX
DT   11-JUN-1999  (first entry)
XX
DE   Lily LAT1 nucleotide sequence.
XX
KW   Lily; Lilium longiflorum; LAT1; qrtH2A; qrtH3; pollen; mature male germline
KW   promoter; sterile; transposase; sperm cell; pollination; seedless fruit;
KW   genetic manipulation; gamete; transposon tagging.
XX
OS   Lilium longiflorum.
XX
FH   Key             Location/Qualifiers
FT   CDS             82..468
FT                   /*tag= a
FT                   /product= "LAT1 protein"
XX
PN   WO9905281-A1.
XX
PD   04-FEB-1999.
XX
PF   24-JUL-1998;    98WO-AU000587.
XX
PR   31-JUL-1997;    97AU-PO8184.
PR   25-JUL-1997;    97AU-PO8234.
XX
PA   (UYME-) UNIV MELB YUNH.
XX
PI   Dolan L, Finch M, Yu J, Xu H;
XX
DR   WPI; 1999-142936/12.
DR   P-PSDB; AAY04129.
XX
PT   New isolated nucleic acid for, e.g. generating male sterile plants
PT   comprises a nucleotide sequence that allows expression of a
PT   transposase gene in generative and sperm cells of a plant
XX
PS   Claim 5; Fig 1; 64pp; English.
XX
CC   The present sequence encodes the LAT1 protein from lily (Lilium
CC   longiflorum). The sequence specifically allows expression of the
CC   gene in generative and sperm cells of a plant. Also described are: (1) a
CC   genetic construct comprising a generative cell and sperm cell specific
CC   promoter operably linked to a transposase, which is capable of inducing
```

```
CC   transposition of a transposable element such that upon expression of the
CC   promoter, the transposase gene is expressed facilitating transposition of
CC   the element; (2) inducing male sterility in a plant by operably linking a
CC   cytotoxic nucleic acid to a promoter, which directs male gamete specific
CC   expression to produce a product from the cytotoxic nucleic acid that
CC   inactivates or kills male gametes in the plant, and (3) a male sterile
CC   plant produced by the method. The male gamete promoter and/or gene is
CC   useful for facilitating male gamete specific transposon tagging, where
CC   the product of pollen grains is screened for a range of phenotypes of
CC   interest, and the transposon tagged plants are used to clone certain
CC   genes. The method is useful for producing seedless fruit or fruit with
CC   reduced seed content, particularly where pollination stimulates fruit
CC   development, and where the lack of fertilisation results in seedless
CC   fruit. The nucleic acids enable specific genetic manipulation of the
CC   male germ line by generating male sterile plants, and facilitating male
CC   gamete specific transposon tagging.
XX
SQ   Sequence 625 BP; 218 A; 117 C; 109 G; 181 T; 0 other;

  Query Match             9.9%;  Score 83.6;  DB 20;  Length 625;
  Best Local Similarity   87.1%;  Pred. No. 6.3e-11;
  Matches  115;  Conservative    0;  Mismatches   14;  Indels    3;  Gaps    2;

Qy        817 CCATCCCATCAACAGAAGTTTTAAGTGGAAACCCATTCCAATAGAGAAGATCGAACAAAG 876
              |||||||||||||||||| ||||||||||| ||||| || |||| |||||| ||||||
Db          2 CCATCCCATCAACAGAAGGTTTAAGTGGAAATCCATTTCATTAGAAAAGATCGGACAAAG 61

Qy        877 GGTA  TTTAAACATACAA-ATGGGGGCAGTGGTGTTTCTTTTTGCTTGCGTTCTCTTCT 933
              ||||  |||| |||||| ||| |||| |||| | || | ||||||||||||||||||
Db         62 GGTACTCTTAAGCATACAACATGAGGGCGGTCGCCCTTTTCTTTGCTTGCGTTCTCTTCT 121

Qy        934 GTATGGTTCACA 945
              ||||||||||||
Db        122 GTATGGTTCACA 133

RESULT 3
AAS46530
ID   AAS46530 standard; DNA; 7814 BP.
XX
AC   AAS46530;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Tumour suppressor gene derived chemically modified sequence #252.
XX
KW   Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW   cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW   cytosine methylation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200168912-A2.
XX
PD   20-SEP-2001.
XX
PF   15-MAR-2001; 2001WO-EP02955.
XX
PR   15-MAR-2000; 2000DE-1013847.
PR   06-APR-2000; 2000DE-1019058.
PR   07-APR-2000; 2000DE-1019173.
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-602752/68.
XX
PT   Fragments of chemically modified genes associated with tumour suppressor
PT   genes and oncogenes, useful in designing primers and probes for
PT   analysing diseases associated with cytosine methylation state e.g.
PT   cancer
XX
PS   Claim 1; SEQ ID No 252; 27pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of 18
CC   bases, of a segment of chemically pretreated DNA (CT DNA) e.g. with
CC   bisulphite, of genes associated with tumour suppression and
CC   oncogenes having a sequence taken from 536 (actually 533 since
CC   numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC   (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC   peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC   form part of a set of probes for detecting the cytosine methylation state
CC   and/or single nucleotide polymorphisms and also to be used in an
CC   array for analysing diseases associated with CpG dinucleotides e.g.
CC   cancers and tumours. The probes can also be used in a method for
CC   ascertaining genetic and/or epigenetic parameters for the diagnosis
CC   and/or therapy of existing diseases or the predisposition to specific
CC   diseases, by analysing cytosine methylations. The parameters may be
CC   compared to another set of genetic and/or epigenetic parameters, the
CC   differences serving as basis for diagnosis and/or prognosis events which
CC   are disadvantageous to patients. The present sequence is one of the
CC   533 genomic sequences derived from tumour suppressor genes and
CC   oncogenes. Sequences with even numbered Seq ID numbers are the
CC   complementary sequence of the corresponding odd numbered sequence (e.g.
CC   ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC   is missing).
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 7814 BP; 1677 A; 101 C; 1779 G; 4257 T; 0 other;

  Query Match             5.8%;  Score 54.6;  DB 22;  Length 7814;
  Best Local Similarity   49.5%;  Pred. No. 0.0015;
  Matches  141;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy         13 AATTAGGTTTGCCTAGGGTTTGCCTAGGTTTAGAGAAATAGTCAAAATTGTCCTATTCTA 72
                | |||  | |    || | | ||         || |   | | || |  | | |   || | ||
Db       3712 ATTTAATTATTTATATGTTTATTAATTAGTTCGTTTATTTGTTATTAGTTTATTAGTTTA 3771

Qy         73 TAGGCATGATTTAGTAGTGAGTTAATTATCCTATAATTTCTCTTCTTGTATGCTGAAATA 132
                ||    |  ||    || | | | |||   ||       | |||| | |  || | | | |  ||
Db       3772 TAATTTTATTTGTTTATTTATTTATGAATTTGTTTATTTATTTATTTATTTGTTTATTTA 3831

Qy        133 ACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTTCTTCCTCAAATTGAGTA 192
                |  || |||| | |      | ||| || | |  |      |||| | |   | | ||   ||
Db       3832 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTCTTTATTAATTTATTTA 3891

Qy        193 TCAACAATTCTTAGATTGCTTTGGTGATTATATTTGATATAATTGTTTGTAAGAATGTGT 252
                | |    |||  |    ||    |||  | |||     ||      | | | | | ||     | | | |
Db       3892 TTATTTATTTATTTTTTTATTTATTTATTTATTTATTTATTAGTTTATTTATTTATTTAT 3951

Qy        253 AGTGAAAAGATTGTGATTATTCATTTCGTTGTTGGACGAATTGTT 297
                       ||   | |  |||||  |||  || ||        | || |
Db       3952 TTATTAATTTATTTATTTATTATTTTATTTATTTATTTATTTATT 3996

RESULT 4
ABL33760/c
ID   ABL33760 standard; DNA; 7851 BP.
XX
AC   ABL33760;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1733.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; ...; eye disease; arteriosclerosis; anaemia;
```

```
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07547.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG ) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising treatment of chemically modified gene, useful
PT   for diagnosis and treatment of diseases associated with abnormal
PT   cytosine methylation -
XX
PS   Claim 1; SEQ ID NO 1733; 92pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention.
XX
SQ   Sequence 7851 BP; 1799 A; 284 C; 2146 G; 3622 T; 0 other;

  Query Match             5.6%;  Score 53;  DB 24;  Length 7851;
  Best Local Similarity   55.8%;  Pred. No. 0.0038;
  Matches  101;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy      97 ATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAA 156
           | ||| ||||  | | || || ||| | |||| | | |    |||| ||||
Db     914 AATATAATATATAATTTATTATTATATAATAATATATATTATTATATTAAAATATATAA 855

Qy     157 TTAAGTTTTGTAGCAATTTCTTGCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGG 216
           |||| | | ||  | ||   | | |||| | ||| | ||| ||| ||
Db     854 TTAATATATATAATATTTATATATTAAAATATAATATATATAATATTTATATATTAAAAT 795

Qy     217 TGATTATATTTGATATAATTGTTTGTAAGAATGTGTATTGAAAAGATTGTGATTATTCAT 276
           ||||||| | ||| || | | | ||| | || |  | | | |  |||| ||
Db     794 ATATTATATATAATATATTTATATATTAAAATATATAATATATATAATATATTTATATAT 735

Qy     277 T 277
           |
Db     734 T 734


RESULT 5
ABL43761
ID   ABL43761 standard; DNA; 7851 BP.
XX
AC   ABL43761;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1734.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiatherosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective, anti-HIV, anti-convulsant, ophthalmological,
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
```

```
KW   antidiabetic; ..., ...or; eye disease; atherosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07547.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG ) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising treatment of chemically modified gene, useful
PT   for diagnosis and treatment of diseases associated with abnormal
PT   cytosine methylation -
XX
PS   Claim 1; SEQ ID NO 1734; 92pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention.
XX
SQ   Sequence 7851 BP; 1811 A; 284 C; 2095 G; 3661 T; 0 other;

  Query Match             5.6%;  Score 53;  DB 24;  Length 7851;
  Best Local Similarity   55.8%;  Pred. No. 0.0038;
  Matches  101;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy      97 ATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAA 156
           | ||| ||||  | | || || ||| | |||| | | |    |||| ||||
Db   6938 AATATAATATATAATTTATTATTATATAATAATATATATTATTATATTAAAATATATAA 6997

Qy     157 TTAAGTTTTGTAGCAATTTCTTGCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGG 216
           |||| | | ||  | ||   | | |||| | ||| | ||| ||| ||
Db   6998 TTAATATATATAATATTTATATATTAAAATATAATATATATAATATTTATATATTAAAAT 7057

Qy     217 TGATTATATTTGATATAATTGTTTGTAAGAATGTGTATTGAAAAGATTGTGATTATTCAT 276
           ||||||| | ||| || | | | ||| | || |  | | | |  |||| ||
Db   7058 ATATTATATATAATATATTTATATATTAAAATATATAATATATATAATATATTTATATAT 7117

Qy     277 T 277
           |
Db   7118 T 7118


RESULT 6
ABQ67111
ID   ABQ67111 standard; DNA; 6673 BP.
XX
AC   ABQ67111;
XX
DT   28-AUG-2002  (first entry)
XX
DE   Human angiogenesis associated polynucleotide SEQ ID NO 141.
XX
KW   Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW   inflammation; rheumatoid arthritis; diabetic retinopathy; ...-clerosis;
KW   macular degeneration; inflammatory bowel disease; Crohn's disease;
```

```
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiarteriosclerotic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200246464-A2.
XX
PD   13-JUN-2002.
XX
PF   06-DEC-2001; 2001WO-EP14320.
XX
PR   06-DEC-2000; 2000DE-1061438.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Schatz H;
XX
DR   WPI; 2002-500450/53.
XX
PI   New nucleic acid fragments from chemically treated
PI   angiogenesis associated genes, useful for determining methylation
PI   status, e.g. in diagnosis or treatment of cancer
XX
PS   Claim 1; SEQ ID NO 141; 41pp; Sequence Listing; German.
XX
CC   The invention relates to a nucleic acid (I) comprising a segment of 18
CC   bases of chemically pretreated DNA of angiogenesis associated genes (II)
CC   having sequences (AB967671-AB967776) or their complements (I), also
CC   related oligomers, are used to evaluate the methylation status and/or
CC   single nucleotide polymorphisms, in angiogenesis-related genes, for
CC   diagnosis and treatment of eye diseases, proliferative retinopathy,
CC   neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
CC   diabetic retinopathy, macular degeneration caused by neovascularisation,
CC   psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC   Crohn's disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 6126 BP; 1641 A; 75 C; 1770 G; 3187 T; 53 other;

  Query Match             5.5%;  Score 52.2;  DB 24;  Length 6673;
  Best Local Similarity   49.8%;  Pred. No. 0.0057;
  Matches  132;  Conservative    0;  Mismatches  133;  Indels    0;  Gaps    0;

Qy      37 TAGGTTTAGAGAAATAGTTAAAATTGTGCTATTCTATAGGCATGATTTAGTAGTGAGTTA 96
           ||   |||||||||   ||   || |  | ||  |   | ||  |||| |||    ||
Db     974 TATTGTTAGAGAAAATATGGTTTTATATTTAGAATAGTATTGTTTAATAGAATTTT 1033

Qy      97 ATTATTTATAATTTCTTTTTGTATGCTTCAAATAACTGGTTCTTTAATGAATAGATAA 156
           ||   |  || | | ||  |    |    |  | | ||| || |   | |
Db    1034 TGTAATGTTGGATATGTTTTTTGTTTGTATTGTTGACTATGGTAGTTAGTGTAATAAGGA 1093

Qy     157 TTAAGTTTTGTAGCAATTTCTTCCTCAAATTGACTATCAACAATTGTTAGATTGCTTTGG 216
           ||||  |    | | ||  |  |||| | | | || | | ||  |  | |     |||
Db    1094 AATGATTTAAAATTTTATTTTATTTAATTAAATGTAAATAGTTATATGTTGTTAGTGG 1153

Qy     217 TGATTATATTTNATATAAATTNTTGTAATAATTTTTATTAAAAAATTTTTATTATTGA 276
           | ||||| | ||||   | |||  |  |  | |||  |  |  | | | |  |
Db    1154 TTATTATTTGGATACGTGGTTTGAGTATAGAGTAGGGGTTAGGTGGTAGTTTTGGT 1213

Qy     277 TTGGTTGTGGCACGAATTGTTAGAG 301
           ||  | ||   |  | |  ||||||
Db    1214 TTTTGGTATTAGAATTATTTACAG 1238


RESULT 7
AAS46573
ID   AAS46573 standard; DNA; 6126 BP.
XX
AC   AAS46573;
XX
DT   18-DEC-2001  (first entry)

XX
DE   Tumour suppressor gene derived chemically modified sequence #295.
XX
KW   Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW   cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW   cytosine methylation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200168912-A2.
XX
PD   20-SEP-2001.
XX
PF   15-MAR-2001; 2001WO-EP02955.
XX
PR   15-MAR-2000; 2000DE-1013847.
PR   06-APR-2000; 2000DE-1019058.
PR   07-APR-2000; 2000DE-1019173.
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2001-602752/68.
XX
PT   Fragments of DNA useful for detecting and treating cancer associated with tumour suppressor
PT   genes and oncogenes, useful in designing primers and probes for
PT   analysing diseases associated with cytosine methylation state e.g.
PT   cancer.
XX
PS   Claim 1; SEQ ID NO 295; 27pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of 18
CC   bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC   bisulphite, of genes associated with tumour suppression and
CC   [illegible] sequences taken from 576 (actually, 573 since
CC   numbers 408, 456 and 500 are missing from the sequence listing) sequences
CC   (IS) and sequences complementary to (IS). The nucleic acid may be a
CC   peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC   form part of a set of probes for detecting the cytosine methylation state
CC   and/or single nucleotide polymorphisms and also to be used in an
CC   array for analysing diseases associated with CpG dinucleotides e.g.
CC   cancers and tumours. The probes can also be used in a method for
CC   ascertaining genetic and/or epigenetic parameters for the diagnosis
CC   and/or therapy of existing diseases or the predisposition to specific
CC   diseases, by analysing cytosine methylations. The parameters may be
CC   compared to another set of genetic and/or epigenetic parameters, the
CC   differences serving as basis for diagnosis and/or prognosis events which
CC   are disadvantageous to patients. The present sequence is one of the
CC   573 genomic sequences derived from tumour suppressor genes and
CC   oncogenes.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 6126 BP; 1753 A; 88 C; 1284 G; 3002 T; 0 other;

  Query Match             5.4%;  Score 51;  DB 22;  Length 6126;
  Best Local Similarity   49.4%;  Pred. No. 0.011;
  Matches  132;  Conservative    0;  Mismatches  135;  Indels    0;  Gaps    0;

Qy      84 TAGTAGTGAGTTAAATATTTATAAATTTTCTTTTTGTATGCTTCAAATAACTGGTTCTTT 143
           |||| | | ||| || |  | ||||   | |  |  ||  | || || |||  |||
Db    1455 TAGTAATTATTTATTTTTTAGTTTAATAGGTTTAGGTAAATTTTTAAATAATTGGTGATTT 1514

Qy     144 AATGAATAGATAATTAAGTTTTGTAGCAATTTCTTCCTCAAATTGACTATCAACAATTGTT 203
           || |  | | ||    ||   || || | |  |  | |  |  |  |  |  |  | |
Db    1515 TATTAGAAAGTGATGTTATTAGAAAGTGATATTTTTTTTATTATAAAAGTGAGTTACTTA 1574

Qy     204 TAGATTGCTTTGGTGATTATATTTNATATAAATTNTTGTAATAATTTTTATTAAAAAGAT 263
```

```
          | | |  | |    ||  | ||  |||| ||| ||  |||     |  |  | | ||
Db   1575 GATAATTTTATTTTTATTTTTTTGTATATTATTTTTAAAAGATATATTTATATATATAT 1634

Qy    264 TGTGATTATTCATTT GTTGTGGACGAATTGTTAGAGCCCCATCGCTAATGCGTTTATAG 323
          | | |    | ||| | || |  |  ||     | ||          |||    | |||
Db   1635 TTTAAAATATAGTTTAGATGATTTATAAAGGTAAATAGTTTGTGTTTTAAATATTAATAT 1694

Qy    324 TACTCGAAATATGTTGGGAATAGAAGA 350
            | | | |||    | ||||
Db   1695 GGTTAAAATTGTGTTTTTTAAGGAAGA 1721


RESULT 8
ABL33840
ID   ABL33840 standard; DNA; 6126 BP.
XX
AC   ABL33840;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1803.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiatherosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07547.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG ) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful
PT   for diagnosis and treatment of diseases associated with abnormal
PT   cytosine methylation.
XX
PS   Claim 1; SEQ ID NO 1803; 42pp - Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention.
XX
SQ   Sequence 6126 BP; 1754 A; 884 C; 1284 G; 3002 T; 0 other;

  Query Match             5.4%;  Score 51;  DB 24;  Length 6126;
  Best Local Similarity   49.4%;  Prod. No. 0.011;
  Matches  142;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy      84 TATTATGTAGTTAATTATTTTTTTTTTTTGTAGTTTTAAATAATTTTTTTTTTTTTTTT 143
           |||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   1459 TAGTAATTATTTATTTTTTAGTTTATAGTGTTTAGGTAAATTTTTAAATAATTGTTGATTT 1518
```

```
Qy    144 AATGAATAGATAATTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 203
          || |  | | | ||    ||   ||  || | ||  |  | | ||  | |  | |
Db   1515 TATTAGAAAGTGATGTTATTAGAAAGTGATATTTTTTTTATAGAAAGTGATTAGTTA 1574

Qy    204 TAGATTGTTTTGTGTGATTATATTTGATATAATTTTTTTTTTTTTTTTTTTTTTTTTTAT 263
          | | |  | |  | || | ||  |||| ||| ||  ||||     |  |  | | ||
Db   1575 GATAATTTTATTTTTATTTTTTTGTATATTATTTTTTTAAAAGATATATTTATATATATAT 1634

Qy    264 TGTGATTATTCATTTGTTGTGGACGAATTGTTAGAGCCCCATCGCTAATGCGTTTATAG 323
          | | |    | ||| | || |  |  ||     | ||          |||    | |||
Db   1635 TTTAAAATATAGTTTAGATGATTTATAAAGGTAAATAGTTTGTGTTTTAAATATTAATAT 1694

Qy    324 TACTCGAAATATGTTGGGAATAGAAGA 350
            | | | |||    | ||||
Db   1695 GGTTAAAATTGTGTTTTTTAAGGAAGA 1721


RESULT 9
AAS46435
ID   AAS46435 standard; DNA; 8093 BP.
XX
AC   AAS46435;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Tumour suppressor gene derived chemically modified sequence #157.
XX
KW   Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW   cancer; tumour; CpG dinucleotide; single nucleotide polymorphism; SNP;
KW   cytosine methylation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200168912-A2.
XX
PD   20-SEP-2001.
XX
PF   15-MAR-2001; 2001WO-EP02955.
XX
PR   15-MAR-2000; 2000DE-1013847.
PR   06-APR-2000; 2000DE-1019058.
PR   07-APR-2000; 2000DE-1019174.
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2001-602752/68.
XX
PT   Fragments of chemically modified genes associated with tumour suppressor
PT   genes and oncogenes, useful in designing primers and probes for
PT   analysing diseases associated with cytosine methylation state e.g.
PT   cancer -
XX
PS   Claim 1; SEQ ID NO 157; 27pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of 18
CC   bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC   bisulphite, of genes associated with tumour suppression and
CC   oncogenes having a sequence taken from 536 (actually 533 since
CC   numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC   (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC   peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC   form part of a set of probes for detecting the cytosine methylation state
CC   and/or single nucleotide polymorphisms and also to be used in an
CC   array for analysing diseases associated with CpG dinucleotides e.g.
CC   cancers and tumours. The probes can also be used in a method for
CC   screening a sample for genetic and/or epigenetic parameters for the diagnosis
CC   and/or therapy of existing diseases or the predisposition to specific
CC   diseases, by analysing cytosine methylations. The parameters may be
CC   compared to another set of genetic and/or epigenetic parameters. The
```

```
CC   differences serving as basis for diagnosis and/or prognosis events which
CC   are disadvantageous to patients. The present sequence is one of the
CC   5+3 genomic sequences derived from tumour suppressor genes and
CC   oncogenes.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 8093 BP; 2303 A; 175 C; 1682 G; 3933 T; 0 other;

  Query Match             5.4%;  Score 51;  DB 22;  Length 8093;
  Best Local Similarity   51.5%;  Pred. No. 0.012;
  Matches  117;  Conservative    0;  Mismatches  110;  Indels    0;  Gaps    0;

Qy      56 AAAAATTGTCCTAATTCTATAAGCATGATTTAGTAGTGAGTTAATTATCCTATAATTTCTCT 115
           | | || |  ; || | || |  |  |||  | |  | || || |   | | || |   |
Db    2242 ATATTTTTATTGTTATGTATGGGTTTTTTTTTTGGAAATTTATGAATTTGTTATGTGAGT 2301

Qy     116 TCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTT 175
           | |||||   |    |||  |  | ||||| || | |||||||  |   || ||  ||||
Db    2302 TTTTGTAAATTGGTTTAAAAGTATATTTAAGGAGTTGATAATTTAAGATTATATGAATTA 2361

Qy     176 CTTGCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTCATATAAT 235
           | ||  |     | ||  |    | ||    |||||  |||| ||   ||| |
Db    2362 GAATTTTAATATTTATTAAAATAAGAGTTGAAATTTTTGCTATTTATTTTTAGAATATT 2421

Qy     236 TGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCATTTCGTT 282
           |      ||| ||  | | ||   ||  | ||  |  ||||  ||
Db   2422 TAAAAAATAGATTGTAAATTAATTATATTAATATTAAATTTTTTT 2468


RESULT 10
ABK92236
ID   ABK92236 standard; DNA; 8093 BP.
XX
AC   ABK92236;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Chemically treated DNA repair gene fragment#23.
XX
KW   DNA repair; cytosine methylation; PMS2L1, PMS2L12, PMS2L2, PMS2L5,
KW   PMS2; L4; PMS2L5, PMS2L6, MSM1, MSH2, NUDT1, TDG, INEIL1, KPS4;
KW   MLH1L; FANCB; XRCC8; ataxia telangiectasia; aging; Bloom's syndrome;
KW   Cockayne syndrome; Nijmegen breakage syndrome, Werner syndrome,
KW   immunodeficiency; trichthiodystrophy; Fanconi's anaemia; solid tumour;
KW   cancer; ds.
XX
OS   Unidentified.
XX
PN   WO200181622 A2.
XX
PD   01-NOV-2001.
XX
PF   06-APR-2001; 2001WO-EP03972.
XX
PR   06-APR-2000; 2000DE-1019058.
PR   07-APR-2000; 2000DE-1019173.
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1044826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-054446/74
XX
PT   New nucleic acid derived from genes associated with DNA repair, useful
PT   for diagnosis, e.g. of ataxia telangiectasia, by determination of
PT   cytosine methylation
XX
PS   Claim 1; SEQ ID NO 45; 25pp + sequence listing; English.

XX
CC   The invention relates to nucleic acids containing a sequence of at least
CC   18 nucleotides of chemically treated DNA of genes associated with DNA
CC   repair, and their complements. The invention also relates to nucleic
CC   acids comprising at least 18 base pairs of the chemically pretreated DNA
CC   of genes associated with DNA repair selected from PMS2L1, PMS2L12,
CC   PMS2L2, PMS2L3, PMS2L4, PMS2L5, PMS2L6, MSM1, MSH2, NUDT1, TDG, INEIL1,
CC   KPS4, MLH1L, FANCB, or XRCC8. Nucleic acids of the invention and related
CC   oligomers, are useful for diagnosis of diseases associated with gene
CC   repair, specifically, ataxia telangiectasia, aging, Bloom's syndrome,
CC   Cockayne syndrome, Nijmegen breakage syndrome or Werner syndrome,
CC   immunodeficiency, trichthiodystrophy, Fanconi's anaemia, solid tumours
CC   and cancer, particularly by determining status of cytosine methylation
CC   and/or by detecting single-nucleotide polymorphisms. Determination of
CC   individual methylation patterns may allow development of individualised
CC   therapies. The sequences given in records ABL92192 ABL92385 represent
CC   chemically pretreated DNA fragments from genes associated with DNA
CC   repair, and their complements.
CC   Note: The sequence data for this patent is not represented in the
CC   specification, but is based on sequence information supplied by the
CC   European Patent Office.
XX
SQ   Sequence 8093 BP; 2303 A; 175 C; 1682 G; 3933 T; 0 other;

  Query Match             5.4%;  Score 51;  DB 24;  Length 8093;
  Best Local Similarity   51.5%;  Pred. No. 0.012;
  Matches  117;  Conservative    0;  Mismatches  110;  Indels    0;  Gaps    0;

Qy      56 AAAATTGTCCTAATTCTATAAGCATGATTTAGTAGTGAGTTAATTATCCTATAATTTCTCT 115
           | | || |  ; || | || |  |  |||  | |  | || || |   | | || |   |
Db    2242 ATATTTTTATTGTTATGTATGGGTTTTTTTTTTGGAAATTTATGAATTTGTTATGTGAGT 2301

Qy     116 TCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTT 175
           | |||||   |    |||  |  | ||||| || | |||||||  |   || ||  ||||
Db    2302 TTTTGTAAATTGGTTTAAAAGTATATTTAAGGAGTTGATAATTTAAGATTATATGAATTA 2361

Qy     176 CTTGCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTCATATAAT 235
           | ||  |     | ||  |    | ||    |||||  |||| ||   ||| |
Db    2362 GAATTTTAATATTTATTAAAATAAGAGTTGAAATTTTTGCTATTTATTTTTAGAATATT 2421

Qy     236 TGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCATTTCGTT 282
           |      ||| ||  | | ||   ||  | ||  |  ||||  ||
Db   2422 TAAAAAATAGATTGTAAATTAATTATATTAATATTAAATTTTTTT 2468


RESULT 11
ABK83973
ID   ABK83973 standard; DNA; 8093 BP.
XX
AC   ABK83973;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human DNA for staging of Astrocytomas #29.
XX
KW   Human; ds; astrocytoma; cytostatic; staging; cysteine methylation; CpG;
KW   bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry;
KW   matrix assisted laser desorption/ionization mass spectrometry.
XX
OS   Homo sapiens.
XX
PN   WO2002028886-A2.
XX
PD   10-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07538.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1044826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
```

```
XX
DE   WPI; 2002-177877/22.
XX
PT   New chemically modified genomic DNA sequences useful in the
PT   characterisation, classification, differentiation, grading, staging,
PT   treatment and/or diagnosis of astrocytomas or predisposition to
PT   astrocytomas
XX
PS   Claim 1; SEQ ID NO 57; 37pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence (I) of at
CC   least 18 bases in length of a segment of chemically pretreated genomic
CC   DNA which has any one of the sequences of (ABK49351-ABK49682) or its
CC   complement. Also included are an oligonucleotide or peptide nucleic
CC   acid (or set thereof) of at least 9 nucleotides which hybridises to (I),
CC   primers for (I), probes for detecting cytosine methylation or single-
CC   nucleotide polymorphisms (SNP) in (I), an array of oligomers
CC   or peptide nucleic acids for analysing diseases associated with the
CC   methylation states of the CpG dinucleotides of (I). The array is useful
CC   for determining genetic and/or epigenetic parameters, classification,
CC   differentiation, grading, staging, treatment and/or diagnosis of
CC   astrocytomas, or the predisposition to astrocytomas by analysing cytosine
CC   methylations. Involves obtaining a biological sample containing genomic
CC   DNA, extracting the genomic DNA, converting cytosine bases which are
CC   unmethylated at the 5 position, in the genomic DNA sample, to uracil or
CC   another base which is dissimilar to cytosine in terms of hybridisation
CC   behaviour, by chemical treatment and amplifying chemically pretreated
CC   genomic DNA fragments using the array and a polymerase, where the
CC   amplificates carry a detectable label. The method further involves
CC   identifying methylation status of one or more cytosine positions, and
CC   analysing methylation status of the cytosine positions by reference to
CC   one or more data sets. The genomic DNA is chemically treated by using a
CC   bisulphite, hydrogen sulphite or disulphite. The amplification
CC   step amplifies DNA which is of particular interest in astrocytoma or
CC   brain tissue, based on the specific genomic methylation status of brain
CC   tissues, as opposed to background DNA. The amplificates carry a
CC   fluorescent label or radionuclide, optionally, the labels of the
CC   amplificates are detectable molecule fragments having a typical mass
CC   which are detected in a mass spectrometer. The fragments of chemically
CC   pre-treated genomic DNA to be amplified, have a single positive or
CC   negative charge for a better detectability in the mass spectrometer.
CC   Preferably, the amplificates or fragments of the amplificates are
CC   detected by matrix assisted laser desorption/ionization mass spectrometry
CC   (MALDI) or using electron spray mass spectrometry (ESI). The
CC   present sequence is one of the chemically pre-treated reference DNA
CC   samples of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 8093 BP; 2303 A; 1759 C; 1682 G; 3333 T; 0 other;

  Query Match             5.4%;  Score 51;  DB 24;  Length 8093;
  Best Local Similarity   51.5%;  Pred. No. 0.012;
  Matches  117;  Conservative    0;  Mismatches  110;  Indels    0;  Gaps    0;

Qy      56 AAAAATCGTCTATTCTATAGGTATGATTTACTATGTAGTTAATTATCCTATAATTTCTCT 115
           |||| | |||  || | | ||| | | | | | | | | | | | | | | | | | |
Db    2242 ATATTTTATATCTTATTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2301

Qy     116 TCTTGTATTTCTAAATAATCTATTTTTTAATCAATATGATATAATTAAGTTTTTAATAATTT 175
           |||||  |  |  ||  | | |||| | | | |||||  |  |||| |  |||
Db    2302 TTTTGTAAATTTGTTTAAAATTATATTTAATTGAGTTGATAATTAATGATATATAATTTA 2361

Qy     176 CTTCCTCAAATTGAGTATCAACAATTGTTAGATTCTTTGTTGATTATATTTGATATAAAT 235
           ||| |  |  |  |  |  |  |||  |  |||||  |||| |||  |||  |  |
Db    2362 CAATTTTAATATTTTATTAAAAATAACATTGAAATTTTTGTTATTATTTTTATAAATATT 2421

Qy     236 TGTTTTAAAGAATTGTATTAAAAAGATTGTGATTATTCATTTCTTT 282
           |  |  |  ||| ||  | |  |  |  | | | | |||  ||
Db    2422 TAAAAAATATAATTGTAAATTTAATTTATATTAATATCAAATTTTTTT 2468
```

```
RESULT 12
ABL49351
ID   ABL49351 standard; DNA; 8093 BP.
XX
AC   ABL49351;
XX
DT   01-MAY-2002  (first entry)
XX
DE   Human MLH1 gene associated with DNA replication SEQ ID NO 51.
XX
KW   Human; cytostatic; neuroprotective; nootropic; immunosuppressant;
KW   gene therapy; gene isolation; DNA replication; CHEK; DNA ic; ATR; CHK1;
KW   ERCC5; SNRPA1; RAD50; L1D2; cytosine methylation; Ataxia telangiectasia;
KW   ATR-X; Bloom's syndrome; tumour; cancer; methylation; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200177377-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-EP04071.
XX
PR   06-APR-2000; 2000DE-1016806.
PR   07-APR-2000; 2000DE-1017124.
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG ) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-017471/02.
XX
PT   New nucleic acid sequences from chemically modified genes associated
PT   with DNA replication, useful for analysing cytosine methylations for
PT   diagnosis and therapy of diseases e.g. Ataxia telangiectasia
XX
PS   Claim 1; SEQ ID NO 51; 23pp + Sequence listing; English.
XX
CC   The invention relates to nucleic acid sequences comprising at least 18
CC   bases of a chemically pretreated gene associated with gene replication,
CC   selected from 94 genes (ABL49301-ABL49394) and/or complementary sequences
CC   associated with DNA replication, CENPB, DNA2L, ATR, CHEK1, ERCC5, SNRPA1,
CC   RAD50 and L1D2. The chemical pretreatment converts cytosine bases
CC   unmethylated at the 5 position to uracil or another base with
CC   hybridisation behaviour dissimilar to cytosine, to enable analysis of
CC   cytosine methylations. The DNA sequences and method are useful in the
CC   diagnosis of diseases (or predisposition to diseases) associated with DNA
CC   replication and in therapy of such diseases, by enabling analysis of the
CC   cytosine methylation patterns of such genes. They are especially useful
CC   in diagnosis and therapy of e.g. Ataxia telangiectasia, ATR X, Bloom's
CC   syndrome, solid tumours and cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification and was supposed to be available directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences. However, the sequence data did
CC   not correspond to that referred to in the specification. The present data
CC   is taken from EPO data for the patent.
XX
SQ   Sequence 8093 BP; 2303 A; 1759 C; 1682 G; 3333 T; 0 other;

  Query Match             5.4%;  Score 51;  DB 24;  Length 8093;
  Best Local Similarity   51.5%;  Pred. No. 0.012;
  Matches  117;  Conservative    0;  Mismatches  110;  Indels    0;  Gaps    0;

Qy      56 AAAAATCGTCTATTCTATAGGTATGATTTACTATGTAGTTAATTATCCTATAATTTCTCT 115
           |||| | |||  || | | ||| | | | | | | | | | | | | | | | | | |
Db    2242 ATATTTTATATCTTATTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2301

Qy     116 TCTTGTATTTCTAAATAATCTATTTTTTAATCAATATGATATAATTAAGTTTTTAATAATTT 175
           |||||  |  |  ||  | | |||| | | | |||||  |  |||| |  |||
Db    2302 TTTTGTAAATTTGTTTAAAATTATATTTAATTGAGTTGATAATTAATGATATATAATTTA 2361
```

```
Qy     176 CTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTGATATAAT 235
           |  ||   |     | ||   |    | ||    ||||||  |||| |||   ||| |
Db    2362 CAATTTTAATATTTTATTAAAATAAGAGTTGAAATTTTGGTATTTATTTTTAGAATATT 2421

Qy     236 TGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCATTTCGTT 282
           |       ||| |||  | | ||   ||  | ||  |  |||| ||
Db    2422 TAAAAAATAGATTGTAAATTTAATTTATATTAATATTAAATTTTTTT 2468


RESULT 13
AAS45388
ID   AAS45388 standard; DNA; 15732 BP.
XX
AC   AAS45388;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Chemically pretreated genomic DNA associated with cell cycle #47.
XX
KW   Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
KW   human immunodeficiency virus; neurodegenerative disorder; solid tumour;
KW   graft versus host disease; glomerular disease; Lewy body disease; cancer;
KW   arthritis; arteriosclerosis; anti HIV; neuroprotective; antiarthritic;
KW   immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
KW   PCR primer.
XX
OS   Homo sapiens.
XX
PN   WO200168911-A2
XX
PD   20-SEP-2001.
XX
PF   15-MAR-2001; 2001WO-EP02945.
XX
PR   15-MAR-2000; 2000DE-1013847.
PR   06-APR-2000; 2000DE-1019058.
PR   07-APR-2000; 2000DE-1019173.
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2001-602731/68
XX
PT   Designing primers and probes for analysing diseases associated with
PT   cytosine methylation state e.g. arthritis, cancer, aging,
PT   arteriosclerosis comprising fragments of chemically modified genes
PT   associated with cell cycle -
XX
PS   Claim 1; SEQ ID No 93; 28pp; English.
XX
CC   Sequences AAS45296-AAS45520 represent chemicall, pretreated genomic DNA
CC   molecules associated with the cell cycle and specific PCR primers of the
CC   invention. The sequences are useful for detecting the methylation state
CC   of all CpG dinucleotides in a sequence and therefore for analysing
CC   associated diseases. By analysing cytosine methylations in the pretreated
CC   DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
CC   of existing diseases or the predisposition to specific diseases can be
CC   ascertained. The parameters may be compared to another set of genetic
CC   and/or epigenetic parameters, the differences serving as basis for
CC   diagnosis and/or prognosis events which are disadvantageous to patients.
CC   The sequences of the invention are useful for the diagnosis and therapy
CC   of HIV infection, neurodegenerative disorders, graft versus host disease,
CC   aging, glomerular disease, Lewy body disease, arthritis,
CC   arteriosclerosis, solid tumours and cancers.
XX
SQ   Sequence 15732 BP; 4638 A; 70 C; 2672 G; 8352 T; 0 other;

  Query Match             5.4%;  Score 50.6;  DB 22;  Length 15732;
  Best Local Similarity   49.0%;  Pred. No. 0.017;
  Matches  164;  Conservative    9;  Mismatches  169;  Indels    2;  Gaps    1;

Qy      81 ATTTAGTAGTGAGTTAATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTC 140
           |||| |   | | ||  || |  |||| ||  |    ||   |  | |||| | |   |
Db    9916 ATTTTGGTTTTAATTTTTTTTTATATATTTGGTAGAATTTATTAGTAAAATTTATTTAGTT 9975

Qy     141 TTTAATGAATAGATAATTAAGTTTTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAAT 200
           |  | |   |          || ||| |     ||  ||       |||| ||   | |||
Db    9976 TGGATTTTTTTTTATTGAGAGGTTTTTGATTTTTGATTTAGTTTTTTGATTAGTTATAAT 10035

Qy     201 TGTTAGATTGCTTTGGTGATTATATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAA 260
           ||     | |||  |  || | | |||| |||    ||   ||  ||| | |   ||
Db   10036 GATTGTTCGGATTTTTTATTTTTTTATGATTTAA--AATTTGTAGGTTGTATGTTTTTAA 10093

Qy     261 GATTGTCATTATTCATTTCGTTGTTCGACGAATTGTTAGACGCCCCATCCCTAATCCCTTA 320
           | | |  ||||| |||| ||  |  | |   | |||       ||   ||||   |||
Db   10094 TAATTTATTTATTTATTTATTTTATTTAGGTTATTTTATTTATTTGATGTATAATTGTTTA 10153

Qy     321 TAGTACTCGAAATATGTTGCGAATACAAGATGAAAAATCCCATTCTTTCGTACTACGAGTA 380
           |||| |    ||| || |  | |   | || ||  |    || || ||  | |   |
Db   10154 AAGTATTTTTTTTATTTTTGTGAAATTTGTTGTAATGTTTTTTTTTTAGTTTTTGATTTT 10213

Qy     381 AAAATTTGTCTTTTCATTATTCCATTGAATGTTAA 415
           |  |||||  |||| || ||   ||  |||||||
Db   10214 ATTATTTCACTTTTTTTTTTTTTTTTATGTTAA 10248


RESULT 14
ABK28233
ID   ABK28233 standard; DNA; 15732 BP.
XX
AC   ABK28233;
XX
DT   23-APR-2002  (first entry)
XX
DE   DNA transcription associated genomic DNA #54.
XX
KW   DNA transcription associated gene; peptide nucleic acid; DNA-oligomer;
KW   DNA, cytosine methylation state, SNP, retroviral infection, gene, ds;
KW   single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
KW   viral infection; Sezary syndrome; haematological disorder; tuberculosis;
KW   immunological disorder; Werner syndrome; developmental disorder;
KW   psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
KW   neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
KW   myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
KW   angiogenesis, [illegible] heart disease, HLA syndrome, gene therapy;
KW   polyglutamine disorder; solid tumour.
XX
OS   Unidentified.
XX
PN   WO200192565-A2.
XX
PD   06-DEC-2001.
XX
PF   06 APR 2001, 2001WO-EP03973.
XX
PR   06-APR-2000; 2000DE-1019058.
PR   07 APR 2000; 2000DE-1019173.
PR   30 JUN 2000; 2000DE-1032529.
PR   01 SEP 2000; 2000DE-1043826.
XX
PA   (EPIG ) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-096946/12.
XX
PT   New nucleic acids or oligomers, useful for diagnosing or treating
PT   diseases associated with DNA transcription, e.g. immunological
PT   disorders, Werner syndrome, psoriasis, myocardial infarction, solid
PT   tumours or cancer -
XX
PS   Claim 1; SEQ ID No 107; 32pp; English.
```

```
XX
CC   The invention relates to a nucleic acid, which comprises a segment of the
CC   chemically pretreated DNA of genes associated with DNA transcription from
CC   one of 846 sequences, and an oligomer, in particular an oligonucleotide
CC   or peptide nucleic acid (PNA) oligomer that hybridises to or is identical
CC   to the chemically pretreated DNA of genes associated with DNA
CC   transcription. The set of oligomer probes are useful for detecting the
CC   cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
CC   in a chemically pretreated genomic DNA. The nucleic acids are useful for
CC   diagnosing or treating diseases associated with DNA transcription
CC   (particularly with the methylation status), e.g. adenosine deaminase
CC   deficiency, viral infection, retrovirus infection, Sezary syndrome,
CC   haemat ological disorders, immunological disorders, Werner syndrome,
CC   tuberculosis, developmental disorders, psoriasis, Kiester's syndrome,
CC   neurological disorders, neurodegenerative disorders, Waardenburg
CC   syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
CC   infarction, hypertension, arteriosclerosis, erythropoiesis, congenital heart
CC   disease, HNK syndrome, arthritis, polyglutamine disorders, solid tumours
CC   or cancer. Sequences ABK28127 ABK28472 represent DNA transcription
CC   associated genomic DNA molecules of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence 15732 BP; 4638 A; 3067 C; 2072 G; 8352 T; 3 other;

  Query Match              5.4%;  Score 50.6;  DB 24;  Length 15732;
  Best Local Similarity   44.0%;  Pred. No. 0.017;
  Matches  264;  Conservative    0;  Mismatches  336;  Indels   27;  Gaps    1;

Qy      81 A-TTAATAATTTATTTAATAATTTATAAATTTTTTTTTGTTATTTTTAAAATAAATTGGTTTT 140
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    9918 ATTTTGTTTTAAATTTTTTTTTATATATTTTGTTATAATTTATTATTAAAAATTATTTATTT 9977

Qy     141 TTTAATTGAATAGTATAATTAATGTTTTTTATGTAATTTTCTTGTTTGAAATTTGTATTATGTAATAAT 205
           | | | |       | | | | |  | | |  | |  | |  |  |||| ||   | |||
Db    9978 TGGATTTTTTTTTTATTGTAGTGGTTTTTGATTTTTGATTTAGTTTTTTGATTAGTTATATAT 10035

Qy     206 TGTTATATTGTTTTTGTTATATATTTGTATATAATTGTTTTTTAATAATGTGTTATTGTAAAA 265
           ||   | |||  | | | | | | ||| |||   ||   ||  | | |  |  |  ||
Db   10036 GATTGTTTGGATTTTTTATTTTTTTATGATTTAA  AATTTTTATGTTTGTTATTTTTTAA 10093

Qy     266 GATTGTTGATTATT ATTTTTTTTTTTTATAATAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTA 325
           | | |  |||| | |   |  | |  |||        ||   ||||  ||  |  |  |||
Db   10094 TAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTATGTATAATTGTTTA 10153

Qy     326 TAGTATTGAAATATGTTGTTAATATAATAGTAAAAATTGTTATTTTTTTGTAGTAGGAGTA 385
           |||| |    ||| | |  | |  | | ||  |    ||  | ||||  | |   |
Db   10154 AATTATTTTTTTTATTTTTGTGAAATTTGTTGTAATTGTTTTTTTTTTTAGTTTTTGATTTT 10213

Qy     386 AAAATTTGTTTTTTATTATTTTTATTGAATTGTTAA 419
           | |||||  |||| ||| ||   ||  |||||||
Db   10214 ATTATTTGAGTTTTTTTTTTTTTTTTTTTATGTTAA 10248


RESULT 15
ABL42543
ID   ABL42543 standard; DNA; 1170 BP.
XX
AC   ABL42543;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID No: 516.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antidementic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW   stroke; ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07547.
XX
PR   30-JUN-2000; 2000DE-1032124.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-143969/17.
XX
PT   Nucleic acid comprising fragment of chemically modified genes, useful
PT   for diagnosis and treatment of diseases associated with abnorma.
PT   cytosine methylation.
XX
PS   Claim 1; SEQ ID NO 516; 32pp; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 11790 BP; 3627 A; 1800 C; 2210 G; 5853 T; 0 other;

  Query Match              5.4%;  Score 50.4;  DB 24;  Length 11790;
  Best Local Similarity   47.0%;  Pred. No. 0.018;
  Matches  156;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy       7 TGTTGGAATTAGGTTTGTTAGGTTTGTTTAGTTTAGAGAAATAGTTTAAAATTGATTTT 66
           | ||  |  || |||| |    || |  |  |  |  |  ||   | | ||||
Db   11194 TTTTTAAtttAGtttttgttttttttttAGtTGGAAATAATGTTTAAATATTATAATTTATTA 11253

Qy      67 ATTGTATAGGTATGATTTAGTATTGAGTTAATTATGTTATAATTTGTGTTGTTGTTATTTT 126
           ||  |  | ||  |||| | | ||| |  | | |||| |  |  |  |||| | |
Db   11254 ATTAGTATTTTTTTGTATTGTATGATTTATTATTGTTTTAAATTTTTTTTTAGTATTTAGAG 11313

Qy     127 TAAATAATTTGTTTTTTTTTTTTAATTAATAATAATTATTTTTTTTATTAAATTTTTTTTAAA 186
           | | | | || ||| | | |  ||| || | |  |  | | | | | | | | | | ||
Db   11314 TTAATAATTATTTTTTTTTTAAAAAAATTAAATAAATTATAATTTTTATTTTTTTTTTAAA 11373

Qy     187 TGAGTATTAATAATTGTTAGATTGTTTTGTTGATTATATTTTGTATAAATTGTTTTTTAATTA 246
           |    || |  ||| | | ||  ||| |  |  |||| |||| | | |||  |
Db   11374 TTTAAATTTAAGATTATAGGGTTTATTTATTTATTAATTTTATATTTGTATTTTTTTTT 11433

Qy     247 ATGTGTATTGTAAAAAGATTGTGATTATTCTATTTGTTGTTTGGATAAATTTTTAGAGGTTTA 306
           | | | | |||| ||| | || || |     || |  | | | | | |  |  |
Db   11434 AAATATGTTTAAAAAATTTGGGTTTTTTTAAAATATTAATTTAAGTATTTGTTTAATTTAT 11493

Qy     307 TTTTAAATTTATATTATTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 366
           |  |||     | | ||  |  | |  |  |
Db   11494 TTTGTAAAATTGTATAATAATAATAATAAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 11545


Search completed: February 24, 2003, 1:36:59
Job time : 344 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 11:59:47 ; Search time 81 Seconds
(without alignments)
3577.896 Million cell updates/sec

Title: US-09-463-480-9
Perfect score: 945
Sequence: 1 gaggggtgttgggaattaggt.....tctcttctgtatgggttcaca 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
3: /cgn2_6/prodata/1/ina/7A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/8A_COMB.seq.*
5: /cgn2_6/prodata/1/ina/backfiles1.seq.*
6: /cgn2_6/prodata/1/ina/backfiles2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42.6	4.5	658	4	US-08-998-416-595	Sequence 595, App
c 2	41.4	4.4	19124	2	US-08-487-926B-13	Sequence 13, Appl
3	40	4.2	1499	4	US-09-122-402B-13	Sequence 13, Appl
4	39.8	4.2	701	4	US-08-998-416-701	Sequence 701, App
5	38.6	4.1	615	4	US-08-998-416-196	Sequence 196, App
6	38.6	4.1	636	4	US-08-998-416-1137	Sequence 1137, Ap
7	38.6	4.1	837	4	US-08-998-416-288	Sequence 288, App
8	38.2	4.0	695	2	US-08-983-795A-36	Sequence 36, Appl
c 9	38.2	4.0	2437	4	US-08-936-165A-191	Sequence 191, App
c 10	37.6	4.0	1979	1	US-08-392-928C-2	Sequence 3, Appli
c 11	37.6	4.0	1979	3	US-09-330-945-3	Sequence 3, Appli
c 12	37.6	4.0	3659	1	US-08-220-151-72	Sequence 72, Appl
c 13	37.6	4.0	3659	1	US-08-303-124-12	Sequence 12, Appl
c 14	37.6	4.0	3659	1	US-08-413-118-72	Sequence 72, Appl
c 15	37.6	4.0	3659	1	US-08-221-657-18	Sequence 18, Appl
c 16	37.6	4.0	3659	2	US-08-184-009-77	Sequence 77, Appl
c 17	37.6	4.0	3659	2	US-08-417-210A-67	Sequence 67, Appl
c 18	37.6	4.0	3659	2	US-08-480-697B-12	Sequence 12, Appl
c 19	37.6	4.0	3659	2	US-08-458-256-77	Sequence 77, Appl
c 20	37.6	4.0	3659	3	US-08-473-446-72	Sequence 72, Appl
c 21	37.6	4.0	3659	4	US-08-460-736-77	Sequence 77, Appl
c 22	37.6	4.0	3659	4	US-09-354-138-48	Sequence 48, Appl
c 23	37.6	4.0	3660	1	US-08-475-063-25	Sequence 25, Appl
c 24	37.6	4.0	3660	1	US-08-207-792-25	Sequence 25, Appl
c 25	37.6	4.0	3661	1	US-08-105-983-221	Sequence 221, App
c 26	37.6	4.0	3661	1	US-08-303-124-7	Sequence 7, Appli
c 27	37.6	4.0	3661	1	US-08-204-729-7	Sequence 7, Appli
c 28	37.6	4.0	3661	1	US-08-709-289-221	Sequence 221, App
c 29	37.6	4.0	3661	1	US-08-458-131-221	Sequence 221, App
c 30	37.6	4.0	3661	2	US-08-480-697B-7	Sequence 7, Appli
c 31	37.2	3.9	7218	1	US-08-232-164-14	Sequence 14, Appl
32	37	3.9	2030	2	US-08-705-947-7	Sequence 7, Appli
c 33	36.8	3.9	4467	1	US-08-565-907A-1	Sequence 1, Appli
c 34	36.8	3.9	4467	2	US-08-910-551B-1	Sequence 1, Appli
c 35	36.8	3.9	4467	2	US-08-909-125A-1	Sequence 1, Appli
36	36.4	3.9	3095	6	5231168-1	Patent No. 5231168
c 37	36.2	3.8	4437	1	US-08-559-302B-72	Sequence 72, Appl
c 38	36.2	3.8	4437	4	US-09-175-928-72	Sequence 72, Appl
c 39	36.2	3.8	4437	4	US-09-798-696-3	Sequence 3, Appli
c 40	36.2	3.8	4613	3	US-08-827-529-27	Sequence 27, Appl
c 41	36.2	3.8	99503	4	US-09-798-696-10	Sequence 10, Appl
42	36	3.8	844	4	US-08-998-416-305	Sequence 305, App
43	36	3.8	1979	1	US-08-592-214A-17	Sequence 17, Appl
44	36	3.8	4479	3	US-09-149-976-17	Sequence 17, Appl
c 45	36	3.8	15225	2	US-08-892-403A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-998-416-595
; Sequence 595, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Gebruchung, Corinne
; TITLE OF INVENTION: [illegible] INVENTION [illegible] ASSAYS [illegible] SYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9626/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: [illegible] A-33376
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919 541-8689
; INFORMATION FOR SEQ ID NO: 595:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 658 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAC1408KP

```
US-08-998-416-595

  Query Match              4.5%;  Score 42.6;  DB 4;  Length 659;
  Best Local Similarity   53.7%;  Pred. No. 0.025;
  Matches  110;  Conservative    0;  Mismatches   94;  Indels    1;  Gaps    1;

Qy      32 TTGGCCAGCTTTAGATAAATATCAAAATTGTGCTATTCTATAGGCATGATTTACTACTG 91
           || ||   | | ||  ||| | |  ||| |  ||| |||     | |||   |  |
Db     186 TTAACTTTATCTTGATTAATATTTATTATTTTGTTATTTTAT TATTTTATTATTTTATT 246

Qy      92 AGTTAATTATGTCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATA 151
           | || ||||  ||| |||  |  |.  |  | .||  |  ||  ||| | | |
Db     247 ATTTTATTATTTTATTATTTTATTATTTTATTATTTATTTATTTATTATTTATTT 306

Qy     152 GATAATTAAGTTTTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGC 211
           || ||| | ||| ||  |||| |  |  | ||  ||| |  ||| |  ||
Db     307 TATTATTTATTTTATTATTTATTTTATTATTTATTTATTTATTATTTATTTATTGTTTG 366

Qy     212 TTTGCTGATTATATTTGATATAATT 236
           |||| |  ||| ||| |  | |||
Db     367 TTTGTTTATTATTTTTTTATTTATT 391


RESULT 2
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
;  GENERAL INFORMATION:
;    APPLICANT:  Sim, Kim L.
;    APPLICANT:  Chitnis, Chetan
;    APPLICANT:  Miller, Louis H.
;    APPLICANT:  Peterson, David S.
;    APPLICANT:  Su, Xin-zhuan
;    APPLICANT:  Wellems, Thomas E.
;    TITLE OF INVENTION:  BINDING DOMAINS FROM PLASMODIUM VIVAX
;    TITLE OF INVENTION:  AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe Martens Olson & Bear
;      STREET:  620 Newport Center Drive  16th Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,826B
;      FILING DATE:  10-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Isralsen, Ned
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH121.001CP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19124 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-487-826B-13

  Query Match              4.4%;  Score 41.4;  DB 2;  Length 19124;
  Best Local Similarity   52.0%;  Pred. No. 0.14;
  Matches   93;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy      66 TATTCTATAGGCATGATTTACTACTGCTGTTAATTATCCTATAATTTCTCTTCTTGTATGC 125
           | || | ||   |  |||  |  | | ||  || |  ||   ||| || . || ||| |
Db   15629 TTTTTTTTATATAAAATTTTTTTTTAATTTTTTTTTGATAATCTTTTCTATTTTTATTC 15570

Qy     126 TCAAATAACTGGTTCTTTAATGAATAGATAATTAATTTTTGTAGTAATTCTTTTTTAAA 185
           |   | ||  |  |||| |  |    | ||||  ||||  |  | |||  ||||
Db   15569 TATCAAAATTTATATTTTATATAATTTTTATATTTTTTTAAAAAAATTTTTCTCCTTTT 15510

Qy     186 TTGAGTATCAATAATTTTTAGTATTTTTGTTCATTATATTTGATATAATTTTGTAA 244
           ||   | |       || || | |   || |  |||||||| || |   | ||| | |
Db   15509 TTTTTTTTTTTTTTTATTTAATAAATTTTTTTTATATTTCATTTTCTTTTTTTCA 15451


RESULT 3
US-09-122-400B-13
; Sequence 13, Application US/09122400B
; Patent No. 6245974
; GENERAL INFORMATION:
;  APPLICANT:  Michalowski, Susan
;  APPLICANT:  Spiker, Steven
;  TITLE OF INVENTION:  MATRIX ATTACHMENT REGIONS
;  FILE REFERENCE:  Michalowski and Spiker
;  CURRENT APPLICATION NUMBER:  US/09/122,400B
;  CURRENT FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  60/054,118
;  PRIOR FILING DATE:  1997-08-06
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 13
;  LENGTH: 1499
;  TYPE: DNA
;  ORGANISM: Nicotiana tabacum
US-09-122-400B-13

  Query Match              4.2%;  Score 40;  DB 4;  Length 1499;
  Best Local Similarity   56.7%;  Pred. No. 0.15;
  Matches   98;  Conservative    0;  Mismatches   75;  Indels    8;  Gaps    1;

Qy       3 AGGGTGTTGGAATTAGGTTTGCCTAGGGTTTGCCTAGGTTTAGATAAATATCTAAAAATTG 62
           | ||||| |  ||| | |   | | |||| | | || |  ||  | ||| ||
Db    1204 ATGGTGTCTGGCATACTTGTTTGGAAATTTGAATCAGGTTAAAATTATGTTGTTAATTTT 1263

Qy      63 TCCTATTCTATAGGCATGATTTAGTACTGA---GTTAATTATCCTATAATTTCTCTTCTTT 119
           | | |  |||||||||  |||     |||    || || | |  ||||   | ||
Db    1264 TATTTTGTTATAGGCATTTTTTGTGCTTGATTGTTTTGTTGTCATTATATTTTATTATTT 1323

Qy     120 GTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAATTTTTGTAGTAATTCTTTTT 179
           | | |   | |||  || ||  ||||  |||||||||| || |||   || ||||
Db    1324 GTAATTTTATTATTATTTTTCTATTAAATTTTAGATAATTAATTTTTATAATTAAATTT 1379


RESULT 4
US-09-998-416-701
; Sequence 701, Application US/09998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Kebischull, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
```

```
;       COUNTRY:  USA
;       ZIP:  27709
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/998,416
;       FILING DATE:  24-DEC-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  CH 3616/96
;       FILING DATE:  31-DEC-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meigs, J. Timothy
;       REGISTRATION NUMBER:  38,241
;       REFERENCE/DOCKET NUMBER:  PA/5 Sandoz A-20174
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  919-541-8587
;       TELEFAX:  919-541-8689
;   INFORMATION FOR SEQ ID NO:  701:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  701 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     ORIGINAL SOURCE:
;       ORGANISM:  PAG1462RP
US-08-998-416-701

  Query Match             4.2%;  Score 39.8;  DB 4;  Length 701;
  Best Local Similarity   49.8%;  Pred. No. 0.14;
  Matches  101;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy      51 TAGTCAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTAATTATCCTATAATT 110
           || | || | | | |  ||  || ||        ||       |  || ||| |  ||||
Db     459 TATTTAATAGTCTACCCTTTAATTGGATATTACTACCTAGTAAATATTTACCTAATAATA 518

Qy     111 TCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGC 170
           | |  |   | || || ||||       | | | ||  | ||   || || |  |  | |
Db     519 TATTATTAAGAATACTTAAAICTAATAATTTATTATCTAAACTATATAAATTAATTAAAC 578

Qy     171 AATTTCTTCCTCAAATTGACTATCAACAATTGTTAGATTGCTTTGCTGATTATATTTGAT 230
           |||  |  | |  |  | | | |  |||  || |||   ||  |  | |  || |
Db     579 TTTTTTATTATTATTTAAATTATTATTAATTAGTAAATTATATTTATTTATTTTATTAAC 638

Qy     231 ATAATTGTTTGTAAGAATGTGTA 253
           ||||| |||| |  |   | ||
Db     639 ATAATTTTTTGGATAATAATATA 661


RESULT 5
US-08-998-416-186
; Sequence 186, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;       COUNTRY:  USA
;       ZIP:  27709
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/998,416
;       FILING DATE:  24-DEC-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  CH 3616/96
;       FILING DATE:  31-DEC-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meigs, J. Timothy
;       REGISTRATION NUMBER:  38,241
;       REFERENCE/DOCKET NUMBER:  PA/5 Sandoz A-20174
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  919-541-8587
;       TELEFAX:  919-541-8689
;   INFORMATION FOR SEQ ID NO:  186:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  615 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     ORIGINAL SOURCE:
;       ORGANISM:  PAG1074RP
US-08-998-416-186

  Query Match             4.1%;  Score 38.6;  DB 4;  Length 615;
  Best Local Similarity   52.1%;  Pred. No. 0.28;
  Matches  136;  Conservative    0;  Mismatches  117;  Indels    7;  Gaps    2;

Qy      37 TAGGTTTAGAGAAATAGTCAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTA 96
           ||  || |   | | ||   ||  |   || || |||    ||  || |||| | | |||
Db      15 TAATTTAATTTATAAAGATTAATATAAACTTTTTTATTATAATATTTAAGTATTAAATTA 74

Qy      97 ATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAA 156
           |||  |||| |||    || ||  ||      ||| | | | |  |||| | |   ||
Db      75 TTTAAACTATTATTATCATTATTTAATAAATTAATTATTTGATTATTAAT-ACTTATTAT 133

Qy     157 TTAAGTTTTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGG 216
           ||| | || ||  | || |||     |||| |  || |  |||  ||| ||   | |
Db     134 ATAATTATTATATAATTTACTT-----AATTCATCATTATTAATATTTATATAATTATAA 188

Qy     217 TGATTATATTTGATATAATTGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCAT 276
           || |||||| |||| | |  |  | ||  | |||     |     |   | ||||| |
Db     189 AAATAATATTTAATATGAATACTATTTAGTCTATGTTCAAATTTTAAATTAGTTATTAAA 248

Qy     277 TTCGTTGTTCGACGAATTGTT 297
           |  |  | |     ||| ||
Db     249 ATATTATTAGATATTATTATT 269


RESULT 6
US-08-998-416-1137
; Sequence 1137, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  No. 6239264arris Corporation
;       STREET:  4054 Cornwallis Road
;       CITY:  Research Triangle Park
;       STATE:  No. 6239264th Carolina
;       COUNTRY:  USA
;       ZIP:  27709
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/998,416
;       FILING DATE:  24 DEC 1997
;       CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/034,207
;       FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;       NAME:  Meigs, J. Timothy
;       REGISTRATION NUMBER:  38,241
;       REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  919 541 9587
;       TELEFAX:  919 541 8689
;  INFORMATION FOR SEQ ID NO:  287:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  838 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;       ORGANISM:  PAG1692RP
US-08-998-416-1147

  Query Match             4.1%;  Score 48.6;  DB 4;  Length 646;
  Best Local Similarity   52.1%;  Pred. No. 0.28;
  Matches  136;  Conservative    0;  Mismatches  119;  Indels    6;  Gaps    2;

Qy      87 TAGGTTTAGAGAAATAGTCAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTA 96
           || || |   | | ||   || |   || || |||    .| || |||| | | |||
Db      15 TAATTTAATTTATAAAGATTAATATAAACTTTTTATTATAATATTTAAGTATTAAATTA 74

Qy      97 ATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAA 156
           ||| |||| |||    |. || ||     ||| | | | |  ||||| | |   ||
Db      75 TTTAAACTATTATTATCATATTTAATAAATTAATTATTGATTATTAAT ACTTATTAT 133

Qy     157 TTAATTTTTTTCTAATTT... ...AAAA.AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 215
           ||| | || || | | |      |||| | || | ||| ||| || | | |
Db     134 ATAATTATTATATAAATTACTT     AATTCATCATTATTAATATTTATATAATTATAA 188

Qy     216 TGAATAATATTTGATAAAAAAAAAAAAAAAAATTTAGTAAAAAAATTTTTTTTTTTTTT 275
           || ||||| ||| | |   | | ||| | |||   |   |   |  ||||| |
Db     189 AAATAATATTTAATATGAATACTATTTAGTCTATGTTCAAATTTTAAATTAGTTATTAAA 248

Qy     276 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 335
           | | | | |      || |
Db     249 ATATTATTAGATATTATTAT 269


RESULT 7
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Phillipsen, Peter
;    APPLICANT:  Fohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Schleichardt, Corinna
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264arris Corporation
;      STREET:  4054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/034,207
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-9587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  288:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  837 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1241RP
US-08-998-416-288

  Query Match             4.1%;  Score 48.6;  DB 4;  Length 837;
  Best Local Similarity   52.1%;  Pred. No. 0.4;
  Matches  136;  Conservative    0;  Mismatches  119;  Indels    6;  Gaps    2;

Qy      87 TAGGTTTAGAGAAATAGTCAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTA 96
           || || |   | | ||   || |   || || |||    .| || |||| | | |||
Db      15 TAATTTAATTTATAAAGATTAATATAAACTTTTTATTATAATATTTAAGTATTAAATTA 74

Qy      97 ATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAA 156
           ||| |||| |||    |. || ||     ||| | | | |  ||||| | |   ||
Db      75 TTTAAACTATTATTATCATATTTAATAAATTAATTATTGATTATTAAT ACTTATTAT 133

Qy     157 TTAATTTTTTTCTAATTT... ...AAAA.AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 215
           ||| | || || | | |      |||| | || | ||| ||| || | | |
Db     134 ATAATTATTATATAAATTACTT     AATTCATCATTATTAATATTTATATAATTATAA 188

Qy     216 TGAATAATATTTGATAAAAAAAAAAAAAAAAATTTAGTAAAAAAATTTTTTTTTTTTTT 275
           || ||||| ||| | |   | | ||| | |||   |   |   |  ||||| |
Db     189 AAATAATATTTAATATGAATACTATTTAGTCTATGTTCAAATTTTAAATTAGTTATTAAA 248

Qy     276 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 335
           | | | | |      || |
Db     249 ATATTATTAGATATTATTAT 269


RESULT 8
US-08-884-206A-36
; Sequence 36, Application US/08884206A
; Patent No. 5986107
;  GENERAL INFORMATION:
;    APPLICANT:  Selitrenn, Genevieve
;    APPLICANT:  Awaad, Gregor
;    TITLE OF INVENTION:  Recombinant DNA Molecules and Expression
```

```
;    TITLE OF INVENTION:  Vectors for Tissue Plasminogen Activator
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BERESKIN & PARR
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,795A
;      FILING DATE:  27-JUN-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  7841-062
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  665 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      CLONE:  Rh 32
US-08-883-795A-36

  Query Match             4.0%;  Score 38.2;  DB 2;  Length 665;
  Best Local Similarity   48.6%;  Pred. No. 0.36;
  Matches  162;  Conservative    0;  Mismatches  168;  Indels    3;  Gaps    2;

Qy      68 ITCTATAGGCATGATTTAGTAGTGAGTTAATTATCCTATAATTTCTCTTCTTGTATGCTC 127
           || ||     ||  |||| | | |  || |||  |  ||| | |  |  | |  |
Db     219 ITATAATTAAATATTTTATAATTAAAATATTTATAATTAAATATTTTATAATTAAAATGT 278

Qy     128 AAATAACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTTCTTCCTCAAATT 187
             |||| |   | ||| || | || |   || |  |   ||   |||| |  | ||| |
Db     279 TTATAATTAAATATTTTATAATTAAAATGTTTA--TAATTACATATTTTATAATTAAAAT 336

Qy     188 GAGTATCAACAATTGTTACATTCCTTTCCTCATTATATTTGATATAATTCTTTCTAACAA 247
           |  ||| |  |  | ||  ||   |   || ||||| | | ||| ||  |  | | ||
Db     337 GTTTATAATTACATATTTTATAATTAAAATGTTTATAATT-ACATATTTTATAATTAAAA 395

Qy     248 TGTCTAGTCAAAACATTCTCATTATTCATTTCTTTGTTCCACGAATTCTTTAGAGCCCCAT 307
           ||| ||    |   | | || ||| |  |  || |      |  | ||| |    |||
Db     396 IGTTTATAATTACATATTTTATAATTAAAATGTTTATAATTACATATTTTATAATTACAT 455

Qy     308 CGCTAATGCCTTATAGTACTCGAAATATGTTGGGAATAGAAGATGAAAAATCCCATTCTT 367
           | ||    |||        | |||| ||   | || |  ||  | |||  |||  ||
Db     456 ATTTTATAAAGTATTTATAATTACATATTTTATAATTAAAGTATTTATAATTACATATTT 515

Qy     368 IGTAGTAGGAGTAAAAATTTCTCTTTTCATTAT 400
           | || |   ||||  ||   |   |   ||||
Db     516 TATAATTAAAGTATTTATAATTACATATTTTAT 548


RESULT 9
US-08-936-165A-191/c
; Sequence 191, Application US/08936165A
; Patent No. 6348582
;  GENERAL INFORMATION:
;    APPLICANT:  Black, Michael
;    APPLICANT:  Burnham, Martin
;    APPLICANT:  Hodgson, John
;    APPLICANT:  Knowles, David
;    APPLICANT:  Lonetto, Michael
;    APPLICANT:  Nicholas, Richard
;    APPLICANT:  Pratt, Julie
;    APPLICANT:  Reichard, Richard
;    APPLICANT:  Rosenberg, Martin
;    APPLICANT:  Ward, Judith
;    TITLE OF INVENTION:  No. 6348582e1 Prokaryotic Polynucleotides,
;    TITLE OF INVENTION:  Polypeptides and Their Uses
;    NUMBER OF SEQUENCES:  534
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/936,165A
;      FILING DATE:  24-SEP-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/027,022
;      FILING DATE:  24-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gimmi, Edward R
;      REGISTRATION NUMBER:  38,891
;      REFERENCE/DOCKET NUMBER:  P50519
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-4478
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  191:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2437 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-936-165A-191

  Query Match             4.0%;  Score 38.2;  DB 4;  Length 2437;
  Best Local Similarity   52.9%;  Pred. No. 0.53;
  Matches   82;  Conservative    0;  Mismatches   73;  Indels    8;  Gaps    0;

Qy      98 TTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAAT 157
           | |||   |||||  | | | |  ||||    | |  |||| |||  |        ||
Db    1010 TCATCTCCTAATTAAATATGGTTCATGCAGCTAATATCGGTTATTTTTTTGTGCTCAAAG 951

Qy     158 TAACTTTTCTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTACATTCCTTTCCT 217
           || |||| |   ||| ||| | ||    | | |  | | ||||    | | |  ||| |
Db     950 AAATTTTTAATTTAATCTCTCCATCGTCATTAATTGCTATAAATTTCGTGTTCGTGTTGTT 891

Qy     218 CATTATATTTGATATAATTCTTTCTAACAATGTCT 252
           | |  ||| |||||  |  ||| | | |||  ||
Db     890 CAATGAATTCGATATTTTCATTTATTAAAATAGGT 856


RESULT 10
US-08-892-828C-4/c
; Sequence 3, Application US/08892828C
; Patent No. 5795962
;  GENERAL INFORMATION:
;    APPLICANT:  IWANAGA, SADAAKI
;    APPLICANT:  MUTA, TATSUSHI
;    APPLICANT:  SEKI, NORIAKI
```

; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; TITLE OF INVENTION: [illegible]
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: [illegible]
; STATE: [illegible]
; COUNTRY: [illegible]
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: [illegible]
; COMPUTER: [illegible]
; OPERATING SYSTEM: [illegible]
; SOFTWARE: [illegible]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; CLASSIFICATION: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: [illegible]
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible]
; TYPE: [illegible]
; STRANDEDNESS: [illegible]
; TOPOLOGY: [illegible]
; MOLECULE TYPE: [illegible]
; FEATURE:
; NAME/KEY: [illegible]
; LOCATION: [illegible]
US-08-392-828C-8

Query Match [illegible]; Score [illegible]; DB 1; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

[illegible]

RESULT 11
US-09-330-945-2
; Sequence 2, Application US/09330945
; Patent No. 6077946
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, SALAAKI
; APPLICANT: MIHA, TAISUKE
; APPLICANT: SEKI, NOBUAKI
; APPLICANT: [illegible]
; TITLE OF INVENTION: DNA ENCODING [illegible]
; TITLE OF INVENTION: AMELOGENIN [illegible]
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, [illegible]
; ADDRESSEE: [illegible]
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/330,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1079 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 194..1027
US-09-330-945-2

Query Match [illegible]; Score [illegible]; DB 1; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

[illegible]

RESULT 12
US-08-220-151-72
; Sequence 72, Application US/08220151
; Patent No. 5520780
; GENERAL INFORMATION:
; APPLICANT: Barlotti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NOVEL CITRIC AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: CANINE HERPESVIRUS GLYCOPROTEINS AND DNA
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gallis, Morris & Safford
; STREET: 530 Fifth Avenue

```
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/229,151
;       FILING DATE:  30-MAR-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Frommer, William S.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  454110-2440
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840 3333
;       TELEFAX:  (212) 840-0712
;       TELEX:  425066 COHEN
;   INFORMATION FOR SEQ ID NO:  72:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3659 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-229-151-72

  Query Match              4.0%;  Score 47.4;  DB 1;  Length 3659;
  Best Local Similarity   15.5%;  Prod. No. 0.87;
  Matches  171;  Conservative    0;  Mismatches  204;  Indels    1;  Gaps    1;

Qy*     29 GCTTTCGTTATGTTTAGAGAAATAGTCAAAATTGTGCTATCTATATGTATTTATA 86
           | ||  | ||   ||    |  | || ||  || | |      ||   ||| || ||||
Db   2165 GATTCACTTAATTTTTTCAGTTCGTAAACTTTCTCCTCGTAATTCACATAATCTAGTA 2126

Qy      87 CTGACTTAATTATCCTATAATTTCTCTTCTTCTATGCTCAAATAACTCGTTCTTTAATGA 146
           |    | || ||  | | |||||  |  |  |||  | |  | |          || | |
Db   2125 GCATTTACATAATTATTTGATTTCATTATTATTATCTTAATTTTATCATAAATTATATAA 2066

Qy     147 ATATATAATTAACTTTTCTTATAAATTTTTTTTAAATTGATATCTAATATTGTTATA 202
           || ||   | | ||   | |  |  | |  | ||     || | | || | | ||
Db   2065 CTAAATTTGTCATTGCATAACAGGTAATACTATACATCATTTAATTAATAAATGTAAAAT 2006

Qy     203 TGCTTTGGTGATTATATTTGATATAATTGTTTGTAAGAATGTGTACTGAAAACATTGTGA 262
           |        |||         ||| || |||||||  | || | || |   | ||
Db   2005 AAACTAACGATAAATACCGATCTATATTCTTCGTAAGAAAATTTAAT-AATACTGTATGG 1947

Qy     263 TTATTCATTTCGTTGTTGCACGAATTGTTAGAGCCCCATCGTTAATGCCTTATAGTACTC 322
           | ||    ||   ||   ||| ||    |  |  ||  |||      | |  | | |
Db   1946 TAATATTATTCCCCGTATAACGGATGACGATGCTACAATGCTGTATATTATTTATCTTA 1887

Qy     323 GAAATATGTTTTTAATAGAAGATGAAAAATCTCATTCTTTGTAGTATACTACTAAAAATTTC 382
           |||   | | | | | | | ||  |      ||||  ||| |     | ||
Db   1886 ATTGTATTTCTGTATAAATACAGCAACATCAAATTCTTCTTACAACTCTCATAATTCAT 1827

Qy     383 TCTTTTCATTATTCCA 404
           | | | |  |||||
Db   1826 ATATCTTATCTTTCCA 1811


RESULT 13
US-08-463-124-12/c
; Sequence 12, Application US/08304124
; Patent No. 5641490
;  GENERAL INFORMATION:
;     APPLICANT:  Paoletti, Enzo
;     APPLICANT:  Taylor, Jill
;     APPLICANT:  Gottig, Russell
;     TITLE OF INVENTION:  Infectious Bursal Disease Virus
;     TITLE OF INVENTION:  Recombinant Fowlpox Vaccine
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Curtis, Morris & Safford
;       ADDRESSEE:  c/o William S. Frommer
;       STREET:  530 Fifth Avenue
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/304,124
;       FILING DATE:
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/918,311
;       FILING DATE:  21-JUL 1992
;       APPLICATION NUMBER:  US 07/746,254
;       FILING DATE:  26-JUL 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Frommer, William S.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  454110-2440
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840 3333
;       TELEFAX:  (212) 840-0712
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3659 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-463-124-12

  Query Match              4.0%;  Score 47.4;  DB 1;  Length 3659;
  Best Local Similarity   15.5%;  Prod. No. 0.87;
  Matches  171;  Conservative    0;  Mismatches  204;  Indels    1;  Gaps    1;

Qy      29 GCTTTCGTTATGTTTAGAGAAATAGTCAAAATTGTGCTATCTATATGTATTTATA 86
           | ||  | ||   ||    |  | || ||  || | |      ||   ||| || ||||
Db   2165 GATTCACTTAATTTTTTCAGTTCGTAAACTTTCTCCTCGTAATTCACATAATCTAGTA 2126

Qy      87 CTGACTTAATTATCCTATAATTTCTCTTCTTCTATGCTCAAATAACTCGTTCTTTAATGA 146
           |    | || ||  | | |||||  |  |  |||  | |  | |          || | |
Db   2125 GCATTTACATAATTATTTGATTTCATTATTATTATCTTAATTTTATCATAAATTATATAA 2066

Qy     147 ATATATAATTAACTTTTCTTATAAATTTTTTTTAAATTGATATCTAATATTGTTATA 202
           || ||   | | ||   | |  |  | |  | ||     || | | || | | ||
Db   2065 CTAAATTTGTCATTGCATAACAGGTAATACTATACATCATTTAATTAATAAATGTAAAAT 2006

Qy     203 TGCTTTGGTGATTATATTTGATATAATTGTTTGTAAGAATGTGTACTGAAAACATTGTGA 262
           |        |||         ||| || |||||||  | || | || |   | ||
Db   2005 AAACTAACGATAAATACCGATCTATATTCTTCGTAAGAAAATTTAAT-AATACTGTATGG 1947

Qy     263 TTATTCATTTCGTTGTTGCACGAATTGTTAGAGCCCCATCGTTAATGCCTTATAGTACTC 322
           | ||    ||   ||   ||| ||    |  |  ||  |||      | |  | | |
Db   1946 TAATATTATTCCCCGTATAACGGATGACGATGCTACAATGCTGTATATTATTTATCTTA 1887

Qy     323 GAAATATGTTTTTAATAGAAGATGAAAAATCTCATTCTTTGTAGTATACTACTAAAAATTTC 382
           |||   | | | | | | | ||  |      ||||  ||| |     | ||
Db   1886 ATTGTATTTCTGTATAAATACAGCAACATCAAATTCTTCTTACAACTCTCATAATTCAT 1827

Qy     383 TCTTTTCATTATTCCA 404
           | | | |  |||||
Db   1826 ATATCTTATCTTTCCA 1811


RESULT 14
US-08-463-124-72/c
```

```
; Sequence 72, Application US/08413118
; Patent No. 5688920
;  GENERAL INFORMATION:
;     APPLICANT:  PAOLETTI, ENZO
;     APPLICANT:  KIMPTON, KEITH L.
;     TITLE OF INVENTION:  ATTENUATED AVIPOX VIRUS VECTORS
;     TITLE OF INVENTION:  CANINE HERPESVIRUS GB, GC AND GD AND USES THEREFOR
;     NUMBER OF SEQUENCES:  128
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  CURTIS, MORRIS & SAFFORD, P.C.
;       STREET:  530 FIFTH AVENUE, 25TH FLOOR
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  UNITED STATES OF AMERICA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  08/413,118
;       FILING DATE:  29-MAR-1995
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/222,171
;       FILING DATE:  30-MAR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  POPPELL, WILLIAM P.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  154.0-287
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840-3333
;       TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  72:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3959 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-413-118-72

  Query Match             4.0%;  Score 87.6;  DB 1;  Length 3959;
  Best Local Similarity   45.5%;  Pred. No. 0.87;
  Matches  171;  Conservative    0;  Mismatches  204;  Indels    1;  Gaps    1;

Qy      29 GCTTTGCTTAGCTTTACACAAATACTCAAAATTGTCCTATTCTATAGGCATGATTTAGTA 88
           |||  | ||  |||     | | |  ||  || | |       ||   |||  || |||||
Db    2185 GATTCATTTAATTTTTTCCAGTTGTAAATTTCTTCTCCTGTAATTCACATAAATCTAGTA 2126

Qy      89 GTGAGTTAATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGA 148
           |    | || ||  | | :|||  |  |  |||  | |  | |        ||  || |
Db    2125 GCATTTATATAATTATTTATTTCATTATATATACTTAAATTTTATCATAAATTATATAA 2066

Qy     149 ATAGATAATTAAATTTTTGTAGCAATTCTTCTCTAAATTGAGTATCAACAATTGTTAGAT 208
            || ||   | | ||   | ||  :     :  :  |      :  | || || || | ||
Db    2065 CTAAATTTGCATTTCATAACACTTAATATGTATACATTTTAATTAATAAATGTAAAAT 2006

Qy     209 TCTTTTGTGAATAAATTGATAAATTTTTTTAAAATTTCTTATTGAAAAGATTGTCA 268
           |      ||          | | | ||  | | |  | ||   | ||
Db    2005 AAATTAATGTATAAATATTTATTATATTTTTTTAATAAAATTTAAT AATATCTATTT 1947

Qy     269 TTATTCATTTCGTTGTTGGACGAATTGTTAGACCTCATTCTAATTCCTTATATTATCTT 328
           | ||   |||   |  |  | | |  |  |  ||  |  |  | |  | |  |  |
Db    1946 TAATATTATTCTCCCTATAACCGATGACGATGGTTACAATCTGTATATTATTTATCTTA 1887

Qy     329 GAAATATTTTGGGAATAGAAGATAAAAAACCCATTCTTTGAGTAAAATTTCT 388
           ||| |  | |  | | |  |  |  |||||  ||| |    | ||| |
Db    1886 ATTCTATTCTTTATAAATATATCTCAATCATCAAATTCTTCTTAGAACTCTCTATAATTCAT 1827

Qy     389 TCTTTCATTATTCCA 404
           | | ||  |||||

Db    1826 ATATTCTATCTTTCTA 1811


RESULT 2
US-08-224-657-48/c
; Sequence 48, Application US/08224657
; Patent No. 5756102
;  GENERAL INFORMATION:
;     APPLICANT:  Paoletti, Enzo
;     APPLICANT:  Tartaglia, James
;     APPLICANT:  Taylor, Jill
;     TITLE OF INVENTION:  CANINE DISTEMPER VIRUS (CDV)
;     TITLE OF INVENTION:  RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE
;     TITLE OF INVENTION:  RECOMBINANTS
;     NUMBER OF SEQUENCES:  122
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Curtis, Morris & Safford, P.C.
;       STREET:  530 Fifth Avenue
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/224,657
;       FILING DATE:  06-APR-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Frommer, William S.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  4544.0-256
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840-3333
;       TELEFAX:  (212) 840-0712
;       TELEX:  42-5060 "FROMS
;  INFORMATION FOR SEQ ID NO:  48:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3959 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-224-657-48

  Query Match             4.0%;  Score 87.6;  DB 1;  Length 3959;
  Best Local Similarity   45.5%;  Pred. No. 0.87;
  Matches  171;  Conservative    0;  Mismatches  204;  Indels    1;  Gaps    1;

Qy      29 GCTTTGCTTAGCTTTACACAAATACTCAAAATTGTCCTATTCTATAGGCATGATTTAGTA 88
           |||  | ||  |||     | | |  ||  || | |       ||   |||  || |||||
Db    2185 GATTCATTTAATTTTTTCCAGTTGTAAATTTCTTCTCCTGTAATTCACATAAATCTAGTA 2126

Qy      89 GTGAGTTAATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATGA 148
           |    | || ||  | | :|||  |  |  |||  | |  | |        ||  || |
Db    2125 GCATTTATATAATTATTTATTTCATTATATATACTTAAATTTTATCATAAATTATATAA 2066

Qy     149 ATAGATAATTAAATTTTTGTAGCAATTCTTCTCTAAATTGAGTATCAACAATTGTTAGAT 208
            || ||   | | ||   | ||  :     :  :  |      :  | || || || | ||
Db    2065 CTAAATTTGCATTTCATAACACTTAATATGTATACATTTTAATTAATAAATGTAAAAT 2006

Qy     209 TCTTTTGTGAATAAATTGATAAATTTTTTTAAAATTTCTTATTGAAAAGATTGTCA 268
           |      ||          | | | ||  | | |  | ||   | ||
Db    2005 AAATTAATGTATAAATATTTATTATATTTTTTTAATAAAATTTAAT AATATCTATTT 1947

Qy     269 TTATTCATTTCGTTGTTGGACGAATTGTTAGACCTCATTCTAATTCCTTATATTATCTT 328
           | ||   |||   |  |  | | |  |  |  ||  |  |  | |  | |  |  |
Db    1946 TAATATTATTCTCCCTATAACCGATGACGATGGTTACAATCTGTATATTATTTATCTTA 1887

Qy     329 GAAATATTTTGGGAATAGAAGATAAAAAACCCATTCTTTGAGTAAAATTTCT 388
           ||| |  | |  | | |  |  |  |||||  ||| |    | ||| |
Db    1886 ATTCTATTCTTTATAAATATATCTCAATCATCAAATTCTTCTTAGAACTCTCTATAATTCAT 1827

Qy     389 TCTTTCATTATTCCA 404
           | | ||  |||||

Db    1826 ATATTCTATCTTTCTA 1811
```

```
                ||| |  | |  |  | |  || |      |||||  ||| |     | || |
Db   1886 ATTGTATTTCTGTATAAATACAGCAACATCAAATTTCTTCTTAGAACTCTCATAATTCAT 1827

Qy    389 TCTTTTCATTATTCCA 404
              | | ||  |||||
Db   1826 ATATCTTATCTTTCCA 1811


Search completed: February 24, 2003, 14:47:09
Job time : 105 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 24, 2003, 14:25:49 ; Search time 87 Seconds
(without alignments)
6394.569 Million cell updates/sec

Title: US-09-463-480-9
Perfect score: 945
Sequence: 1 qqaqqqtcttggaatlaqqt.....tctcttctqtatqqttcaca 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 442118 seqs, 280819700 residues

Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: [illegible]:*
2: [illegible]:*
3: [illegible]:*
4: [illegible]:*
5: [illegible]:*
6: [illegible]:*
7: [illegible]:*
8: [illegible]:*
9: [illegible]:*
10: [illegible]:*
11: [illegible]:*
12: [illegible]:*
13: [illegible]:*
14: [illegible]:*

Pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41.6	4.4	321	10	US-09-960-352-1267	Sequence 1267, Ap
2	41.6	4.4	405	10	US-09-969-372-873	Sequence 873, App
3	41.6	4.4	335913	9	US-09-754-853A-2	Sequence 2, Appli
4	41.6	4.4	335913	9	US-09-754-853A-3	Sequence 3, Appli
5	41.4	4.4	479	10	US-09-960-352-12872	Sequence 12872, A
c 6	40.4	4.3	3924	12	US-10-014-090-188	Sequence 188, App
7	40	4.2	1499	10	US-09-816-894-13	Sequence 13, Appl
8	39.6	4.2	314	10	US-09-960-352-1036	Sequence 1036, Ap
c 9	39.6	4.2	2090	9	US-09-938-842A-3788	Sequence 3788, Ap
c 10	39	4.1	127197	9	US-09-754-853A-1	Sequence 1, Appli
11	39	4.1	513509	9	US-09-754-853A-4	Sequence 4, Appli
12	38.6	4.1	2524	10	US-09-764-978-403	Sequence 403, App
13	38.6	4.1	640681	10	US-09-790-988-1	Sequence 1, Appli
14	38.4	4.1	1887	10	US-09-970-927A-178	Sequence 178, App
c 15	38.4	4.1	640681	10	US-09-790-988-1	Sequence 1, Appli
c 16	38.2	4.0	2437	10	US-09-939-980-191	Sequence 191, App
17	38	4.0	13865	10	US-09-070-927A-200	Sequence 200, App
c 18	38	4.0	31208	10	US-09-852-067-3	Sequence 3, Appli
19	37.8	4.0	2000	9	US-09-938-842A-2746	Sequence 2746, Ap
20	37.8	4.0	2000	9	US-09-938-842A-4612	Sequence 4612, Ap
21	37.8	4.0	2000	9	US-09-938-842A-5273	Sequence 5273, Ap
c 22	37.6	4.0	208	10	US-09-960-352-4638	Sequence 4638, Ap
23	37.6	4.0	516	10	US-09-960-352-5285	Sequence 5285, Ap
c 24	37.6	4.0	335913	9	US-09-754-853A-2	Sequence 2, Appli
c 25	37.6	4.0	335913	9	US-09-754-853A-3	Sequence 3, Appli
26	37.4	4.0	480	10	US-09-960-352-5301	Sequence 5301, Ap
c 27	37.4	4.0	1565	10	US-09-070-927A-559	Sequence 559, App
c 28	37.2	3.9	416	10	US-09-960-352-4584	Sequence 4584, Ap
29	37	3.9	414	10	US-09-960-352-6528	Sequence 6528, Ap
30	37	3.9	426	10	US-09-960-352-8406	Sequence 8406, Ap
c 31	37	3.9	537	9	US-09-796-692-8190	Sequence 8190, Ap
32	37	3.9	172637	10	US-09-805-458A-3	Sequence 3, Appli
33	36.8	3.9	413	10	US-09-960-352-2919	Sequence 2919, Ap
34	36.8	3.9	426	10	US-09-960-352-6529	Sequence 6529, Ap
c 35	36.8	3.9	2000	9	US-09-938-842A-5412	Sequence 5412, Ap
36	36.4	3.9	401	9	US-09-946-807-408	Sequence 408, App
37	36.4	3.9	401	10	US-09-795-668-408	Sequence 408, App
38	36.4	3.9	401	10	US-09-795-686-408	Sequence 408, App
c 39	36.4	3.9	666	10	US-09-925-299-705	Sequence 705, App
c 40	36.4	3.9	78056	9	US-10-109-551-1	Sequence 1, Appli
41	36.4	3.9	1503841	9	US-09-946-807-1	Sequence 1, Appli
42	36.4	3.9	1503841	10	US-09-795-668-1	Sequence 1, Appli
43	36.4	3.9	1503841	10	US-09-795-686-1	Sequence 1, Appli
c 44	36.2	3.8	4437	10	US-09-753-143-72	Sequence 72, Appl
c 45	36.2	3.8	4437	10	US-09-962-842-112	Sequence 112, App

ALIGNMENTS

RESULT 1
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 1267
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 06 LIB7058 042 Q1 K1-B9
US-09-960-352-1267

Query Match 4.4%; Score 41.6; DB 10; Length 321;
Best Local Similarity 51.0%; Pred. No. 0.35;
Matches 98; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

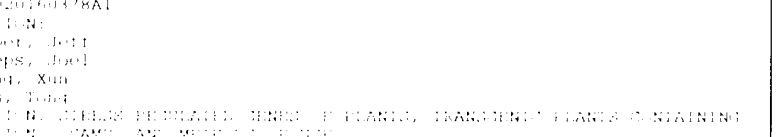
```
Qy      94 TTAATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTGGTTCTTTAATCAATAGA 153
           ||| |||| | | ||    || || |||  | | || | |  || |||     |||
Db      99 TTATTTATTTTTTTTTATATTTTTTTTATTATTATATTATTTTTATTTTTATTATATT 158

Qy     154 TAATTAACTTTTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTT 213
           || || | |||| |   | ||  ||  |   |||   | |     | | ||  ||| ||
Db     159 TATTTAATTTTATTATATTTCTTTATTATTATTTTTTTTATTTTAATTTTTATTTTTT 218

Qy     214 TCCTCATTATATTTGATATAATTCTTTCTAACAATTCTCTATTAAAACATTCTCATTATT 273
           |   ||| | ||  |||  || ||| |   || | || | | |   || | |  ||
Db     219 TTTATATTTTATTTTTATTTTATTTTTTTTTATATTTAATTATATTTTTTTTATTATTT 278

Qy     274 CATTTCGTTCTT 285
           | | |   | ||
Db     279 CTTATTTATTTT 290
```

```
RESULT 2
US-09-969-373-873
; Sequence 873, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Ellerbe, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean cDNAs and Methods of Constructing
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/771,695
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,142
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,760
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 873
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-873

  Query Match             4.4%;  Score 41.6;  DB 10;  Length 405;
  Best Local Similarity   50.5%;  Pred. No. 0.38;
  Matches  101;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy      49 AATAGTCAAAATTGTCCTATTCTATAGGCATGATTTACTAGTGAGTTAATTATCCTATAA 108
           | ||||||| |||||   | | | ||     | |||| || ||| | || ||    ||||
Db     136 ATTAGTCAATATTGTAAAACTTTTACATTAGTTTTATTATTGAATAAAATAAATAATAA 195

Qy     109 TTTCTCTTCTTGTATGTTCAAATAACTGGTTCTTTAATGAATAGTAATTAAGTTTGTA 168
           | |  |   || |   || || | ||   |  ||  ||| |||  ||  | |
Db     196 AAGTTATATTATCATTCCACAAAAATTACTATATATATATATATATATATATATAT 255

Qy     169 GCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTG 228
             | |       | | |   |||    | ||| |  |||  | |  | || || ||
Db     256 ATATATATGGTACAATAATATATATGGTAATTTGCTTAATTATTATAATAATCATTTTAA 315

Qy     229 ATATAATTGTTTGTAAGAAT 248
           | || || |||  || || |
Db     316 AAATGATAGTTATATGATT 335


RESULT 3
US-09-754-853A-2
; Sequence 2, Application US/09754853A
; Publication No. US20040005491A1
; GENERAL INFORMATION:
;  APPLICANT: Hauge, Brian M.
;  APPLICANT:  Parnell, Laurence D.
;  APPLICANT:  Parsons, Jeremy D.
;  APPLICANT:  Wang, Ming Li
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Soybean Cyst Nematode Resistance
;  FILE REFERENCE: 38-10(15810)B
;  CURRENT APPLICATION NUMBER: US/09/754,853A
;  CURRENT FILING DATE:  2001-01-05
;  PRIOR APPLICATION NUMBER: US 60/174,880
;  PRIOR FILING DATE: 2000-01-07
;  NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 2
;   LENGTH: 335913
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (45163)..(45414),(45459)..(45509),(46941)..(48763),(48975)..(49573)
;   OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-2

  Query Match             4.4%;  Score 41.6;  DB 9;  Length 335913;
  Best Local Similarity   50.5%;  Pred. No. 3.7;
  Matches  101;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy      49 AATAGTCAAAATTGTCCTATTCTATAGGCATGATTTACTAGTGAGTTAATTATCCTATAA 108
           | ||||||| |||||   | | | ||     | |||| || ||| | || ||    ||||
Db 194336 ATTAGTCAATATTGTAAAACTTTTACATTAGTTTTATTATTGAATAAAATAAATAATAA 194395

Qy     109 TTTCTCTTCTTGTATGTTCAAATAACTGGTTCTTTAATGAATAGTAATTAAGTTTGTA 168
           | |  |   || |   || || | ||   |  ||  ||| |||  ||  | |
Db 194396 AAGTTATATTATCATTCCACAAAAATTACTATATATATATATATATATATATATAT 194455

Qy     169 GCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTG 228
             | |       | | |   |||    | ||| |  |||  | |  | || || ||
Db 194456 ATATATATGGTACAATAATATATATGGTAATTTGCTTAATTATTATAATAATCATTTTAA 194515

Qy     229 ATATAATTGTTTGTAAGAAT 248
           | || || |||  || || |
Db 194516 AAATGATAGTTATATGATT 194535


RESULT 4
US-09-754-853A-3
; Sequence 3, Application US/09754853A
; Publication No. US20040005491A1
; GENERAL INFORMATION:
;  APPLICANT: Hauge, Brian M.
;  APPLICANT:  Parnell, Laurence D.
;  APPLICANT:  Parsons, Jeremy D.
;  APPLICANT:  Wang, Ming Li
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Soybean Cyst Nematode Resistance
;  FILE REFERENCE: 38-10(15810)B
;  CURRENT APPLICATION NUMBER: US/09/754,853A
;  CURRENT FILING DATE:  2001-01-05
;  PRIOR APPLICATION NUMBER: US 60/174,880
;  PRIOR FILING DATE: 2000-01-07
;  NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 3
;   LENGTH: 335913
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (46798)..(48763),(48975)..(49573)
;   OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-3

  Query Match             4.4%;  Score 41.6;  DB 9;  Length 335913;
  Best Local Similarity   50.5%;  Pred. No. 3.7;
  Matches  101;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy      49 AATAGTCAAAATTGTCCTATTCTATAGGCATGATTTACTAGTGAGTTAATTATCCTATAA 108
           | ||||||| |||||   | | | ||     | |||| || ||| | || ||    ||||
Db 194336 ATTAGTCAATATTGTAAAACTTTTACATTAGTTTTATTATTGAATAAAATAAATAATAA 194395

Qy     109 TTTCTCTTCTTGTATGTTCAAATAACTGGTTCTTTAATGAATAGTAATTAAGTTTGTA 168
           | |  |   || |   || || | ||   |  ||  ||| |||  ||  | |
Db 194396 AAGTTATATTATCATTCCACAAAAATTACTATATATATATATATATATATATATAT 194455

Qy     169 GCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATTGCTTTGGTGATTATATTTG 228
             | |       | | |   |||    | ||| |  |||  | |  | || || ||
Db 194456 ATATATATGGTACAATAATATATATGGTAATTTGCTTAATTATTATAATAATCATTTTAA 194515

Qy     229 ATATAATTGTTTGTAAGAAT 248
           | || || |||  || || |
Db 194516 AAATGATAGTTATATGATT 194535


RESULT 5
US-09-960-452-12872
; Sequence 12872, Application US/09960452
; Patent No. US20020137139A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10299)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12872
;   LENGTH: 479
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 55 LIB3057 018 Q1 K1 F1
US-09-960-352-12872

  Query Match             4.4%;  Score 41.4;  DB 10;  Length 479;
  Best Local Similarity   48.5%;  Pred. No. 0.45;
  Matches  114;  Conservative    0;  Mismatches  121;  Indels    0;  Gaps    0;

Qy      48 AAATAGTCAAAAATTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTAATTATCCTATA 107
           | ||  | ||  || | |  |||| |||     |  |||| |  | | ||  ||   || |
Db     245 ATATTTTTAACTTTTTTTTTATTTTATCTTTTTTTTTTATTTTTTATTTTTTTTATCTTTT 304

Qy     108 ATTTCTCTTCTTGTATGCTCAAATAACTCGTTCTTTAATGAATAGATAATTAAGTTTTCT 167
           |||| | || |  | |  |  ||     || |||      |  || || | |||| |
Db     305 ATTTTTTTTATGTTTTTTTTTTATTTTATTTTTTTTTCATTTTTATTATTTTATTTTTTT 364

Qy     168 AGCAATTTCTTCCTCAAATTGACTATCAACAATTGTTAGATTGCTTTGGTGATTATATTT 227
           | ||| ||  | | ||| | | |    | || ||   ||  ||   |  || | |||
Db     365 TTAATTTTATTTATTATAATTTAATTTATTTATTTTTTTTTTTTTTATATATTTTTTTTT 424

Qy     228 GATATAATTGTTTCTAAGAATGTCTAGTCAAAACATTGTCATTATTCATTTCGTT 282
           | |     | | |     | | |  | |    ||| | ||  ||| ||| ||
Db     425 ATTTTTTAATTATTTTTTTTTTTTTTATTATTTTATTATTATATTTCTTTTATTT 479


RESULT 6
US-10-044-090-188/c
; Sequence 188, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 188
;   LENGTH: 3924
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 085713.1
;   NAME/KEY: unsure
;   LOCATION: 3902, 3912, 3918, 3920
;   OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-188

  Query Match             4.3%;  Score 40.4;  DB 12;  Length 3924;
  Best Local Similarity   56.9%;  Pred. No. 1.6;
  Matches   74;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy      90 TGAGTTAATTATCCTATAATTTCTCTTCTTGTATCCTCAAATAACTCGTTCTTTAATGAA 149
           |||||   ||    | |   ||   ||||||| | ||||  | | || | |   | |
Db    3219 TGAGTAGGTTCTCTTTTCTCTTAAGTTCTTGTGTTCTCATGTTAGTGTTGTTCAGACGCT 3160

Qy     150 TAGATAATTAACTTTTGTAGCAATTTCTTCCTCAAATTGAGTATCAACAATTGTTAGATT 209
           | | |     |||||| |   ||||||| |  |||| | | ||||   || ||| ||
Db    3159 AGGTTTAGGCTTTTTGTTGTTGTTCCTCTTCTTCAATTAAATTTCAAGTTTTTTTAAATA 3100

Qy     210 GCTTTGGTGA 219
           |||||  |||
Db    3099 GCTTTCCTGA 3090


RESULT 7
US-09-816-894-13
; Sequence 13, Application US/09816894
; Patent No. US20020073448A1
; GENERAL INFORMATION:
;  APPLICANT: Michalowski, Susan
;  APPLICANT:  Spiker, Steven
;  TITLE OF INVENTION: MATRIX ATTACHMENT REGIONS
;  FILE REFERENCE: 5681.4001XDV
;  CURRENT APPLICATION NUMBER: US/09/816,894
;  CURRENT FILING DATE:  2001-06-20
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 1499
;   TYPE: DNA
;   ORGANISM: Nicotiana tabacum
US-09-816-894-13

  Query Match             4.2%;  Score 40;  DB 10;  Length 1499;
  Best Local Similarity   55.7%;  Pred. No. 1.5;
  Matches   98;  Conservative    0;  Mismatches   78;  Indels    2;  Gaps    1;

Qy       3 AGGGTGTTGGAATTAGGTTTGGCTAGGGTTTGCCTAGGTTTAGAGAAATAGTCAAAATTG 62
           | |||||  |   ||| | |    | | ||||  |  |  || | |  ||  |  || ||
Db    1204 ATGGTGTCTGGCATAGTTGTTTGGAAGATTGAATCAGGGTaAAAAGTATGTTGTAATTTT 1263

Qy      63 TCCTATTCTATAGGCATGATTTAGTAGTGA---GTTAATTATCCTATAATTTCTCTTCTT 119
           |  | |  |||||||||  |||     |||    ||  || || | |   ||||  |  ||
Db    1264 TATTTTGTTATAGGCATTTTTTGTGCTTGATTGTTTTGTTGTCATTATATTTTATTATTT 1323

Qy     120 GTATGCTCAAATAACTCGTTCTTTAATGAATAGATAATTAAGTTTTCTAGCAATTT 175
           | | |   | |||  || ||  ||||   ||||||||| || |||   || ||||
Db    1324 CGAAGTGTATATATATGTTTGATTAAAATATAGATAATCAATTTTATAAGAAATTT 1379


RESULT 8
US-09-960-352-1036
; Sequence 1036, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 1036
;   LENGTH: 344
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 05 LIB3058 031 Q1 K1 R1
US-09-960-352-1036

  Query Match             4.2%;  Score 39.6;  DB 10;  Length 344;
  Best Local Similarity   47.9%;  Pred. No. 1.1;
  Matches  114;  Conservative    0;  Mismatches  121;  Indels    0;  Gaps    0;

Qy      60 TTGTCCTATTCTATAGGCATGATTTAGTAGTGAGTTAATTATCCTATAATTTCTCTTCTT 119
           || |  | |  | ||    |  |||| |  |   || ||| |  ||| |||| | || ||
Db      61 TTTTTTTTATTTTATCTTTTTTTTATTTATTTTTTATTTTTTTATATTTTTTTTTTTT 120
```

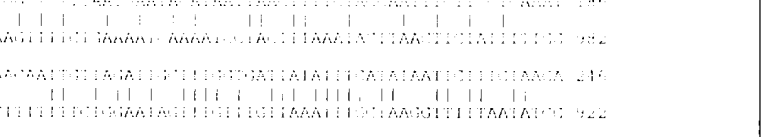
RESULT 9
US 09 938 842A 3788/c
; Sequence 3788, Application US/09 938 842A
; Patent No. US20020078481
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: SIGNAL PEPTIDES DIRECTING ENDOPLASMIC RETICULUM, CHLOROPLAST RETICULUM CONTAINING
; TITLE OF INVENTION: PLANT AND SEED CELLS
; FILE REFERENCE: SCRIPPS.003A
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5479
; SEQ ID NO 3788
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US 09 938 842A 3788

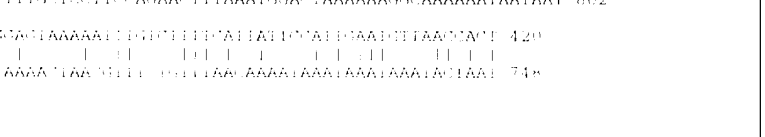
Query Match 4.2%; Score 45.6; DB 9; Length 2000;
Best Local Similarity 45.9%; Pred. No. 2;
Matches 136; Conservative 0; Mismatches 159; Indels 0; Gaps 0;



RESULT 10
US 09 754 854A 1/c
; Sequence 1, Application US/09 754 854A
; Publication No. US20030054918A1
; GENERAL INFORMATION:
; APPLICANT: Baudo, Brian M.
; APPLICANT: Parnell, Lawrence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 98 R(1581b)B
; CURRENT APPLICATION NUMBER: US/09/754,854A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 1
; LENGTH: 127197
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 51[illegible] Region 52
US 09 754 854A-1

Query Match 4.2%; Score 45.6; DB 9; Length 127197;
Best Local Similarity 46.0%; Pred. No. 12;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;





RESULT 11
US 09 754 854A-4
; Sequence 4, Application US/09 754 854A
; Publication No. US20030054918A1
; GENERAL INFORMATION:
; APPLICANT: Baudo, Brian M.
; APPLICANT: Parnell, Lawrence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 98 R(1581b)B
; CURRENT APPLICATION NUMBER: US/09/754,854A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 4
; LENGTH: 513509
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1[illegible])..(1[illegible])
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(513509)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 31B01A region A3
US 09 754 854A-4

```
  Query Match             4.1%;  Score 99;  DB 9;  Length 512509;
  Best Local Similarity   48.4%;  Pred. No. 18;
  Matches  108;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy      56 AAAATTGTGCTATTCTATATAGCTATGATTTAGTAGTGAGTTAATTATTCTTATAATTTCTCT 115
           |.|||     ||| | | |     | || || |          ||| |    | |||| | | |
Db 293882 AAAATGTCAATATAATTTTTCATTTATCTACTTAGTTTATAAATTGAGTTTAATGTTTAT 293941

Qy     116 TCTTGTATGCTCAAATAACTGGTTCTTTAATGAATAGATAATTAAGTTTTGTAGCAATTT 175
           | |  ||| |  || |  |  || | | || | |  || || || | ||||   ||||
Db 293942 ACATTAATGGTATAAAATATTTTATATTATCATTTAATTATAAATTATTGTTGAATTA 294001

Qy     176 CTTCCTCAAATTCACTATCAACAATTGTTACATTCGTTTCGTGATTATATTTGATATAAT 235
           ||   | | | | |   || ||   |||| ||  |    |||| |  | ||   | ||
Db 294002 TTTAAGATAATTATTTAAAAAAAAAATTAGCATGTATAATAGATTGTGATGGAATGATT 294061

Qy     236 TGTTTGTAAGAATGTGTAGTGAAAAGATTGTGATTATTCATTT 278
           || |   | |  || || |   || | |  || || | |
Db 294062 GGGTTTTTTACACGGCTAATGCATGATTTTTTTTTCTCTTAT 294104


RESULT 12
US-09-764-878-403
; Sequence 403, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121
;  CURRENT APPLICATION NUMBER: US/09/764,878
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 403
;   LENGTH: 2534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-878-403

  Query Match             4.1%;  Score 98.6;  DB 10;  Length 2534;
  Best Local Similarity   50.4%;  Pred. No. 3.8;
  Matches   95;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy      83 TTAGTAGTGAGTTAATTATCCTATAATTTCTCTTCTTGTATGCTCAAATAACTCGTTCTT 142
           ||| ||   |||| || |  | || |||   || |   |  || ||| | |  | |  |
Db    1664 TTAATAACAAGTTGTTTGTATTTTATTTTTGTTTATACCAGTCTTAAAGATCCAGGTTCT 1723

Qy     143 TAATCAATAGATAATTAACCTTTTGTAGCAATTTCTTCTCAAATTGAGTTATCAACAATTT 202
           ||| |:   || | |   |  | |  | | |  |           | ||  |  | ||
Db    1724 GAATAAAAAAATTAATTGATACAATTGATGTGTGCTGGGGTTTGGAACTAAAAGTAGTTT 1783

Qy     203 TTAGATTGCTTTTTGTGATTATATTTCATATAATTTGTTTCTAAGAATCTCTTATTCAAAAGA 262
           | | ||  | ||| || | ||||  |||  || ||||:     ||||||| | |  ||
Db    1784 CAACAGTGCGTGGGTTATGACATTTCTTATGTTTCTTTGTTCATGTGTGTATTTAGTAGT 1843

Qy     263 TTTTGATTA 271
           | | | |
Db    1844 TAATTTAA 1852


RESULT 13
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
;  APPLICANT: SHIGENOBU, SHUJI
;  APPLICANT:  WATANABE, HIDEMI
;  APPLICANT:  HATTORI, MASAHIRA
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;  FILE REFERENCE: 081356/0159
;  CURRENT APPLICATION NUMBER: US/09/790,988
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: JP2000-137160
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

  Query Match             4.1%;  Score 98.6;  DB 1;  Length 640681;
  Best Local Similarity   50.2%;  Pred. No. 25;
  Matches  122;  Conservative    0;  Mismatches  121;  Indels    2;  Gaps    1;

Qy     153 ATAATTAAGTTTTGTAGCAATTTCTTCTCAAATTGAGTATCAACAATTGTTACATTCGT 212
           ||  |||| |||  ||          ||    |||||    |  || |   ||| ||| |
Db 124789 ATCGTTAACTTTGATAAATGAAATTTATAAAAATTTTACAGGAAAACAAGTTTTATTAAT 124848

Qy     213 TTCGTGATTATATTTGATATAATTGTTTGTAAGAATCTCTTATTCAAAAGATTTTGATTAT 272
           ||     || | |  |||  || |||||    |||||  |    ||   | ||| |||  |
Db 124849 TTTTAATAAATTTGTTTTATTTTTTTTTTTCTTAATGTGTGTCAGAAATGTAATGACATATT 124908

Qy     273 TTA-TTTSTTTTTSSSSSAATTTTAGATSSSATSSTTAATSSTTATATTATSSA 330
           | |   | |     |  || || |  |\ ||      ||| ||  |   |||  || |
Db 124909 TAAATGTACTAATTGTGATGATTATTTGAAATTAATCAATATGTAAATATTTATTTC 124968

Qy     331 AATATGTTGGGAATAGAAGATGAAAAATSSSATSTTTTTAGTATAATAAAAATTGSC 390
           | || |||   |   | || ||||| || ||| |||   |   ||  | |  | |
Db 124969 AAAT...T...GA...GAAT.AAAAA. .AAT....T.T.TATTGTGCATTTTTTC 125028

Qy     391 TTT 393
           | |
Db 125029 TCT 125031


RESULT 14
US-09-078-927A-178
; Sequence 178, Application US/09078927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles A. Kunsch
;                   Patrick J. Dillon
;                   Steven Barash
;        TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;        NUMBER OF SEQUENCES: 982
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MS-DOS version 6.2
;             SOFTWARE: ASCII text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/078,927A
;             FILING DATE: 04-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/046,655
;             FILING DATE: 1997-05-16
;             APPLICATION NUMBER: 60/044,031
;             FILING DATE: 1997-05-06
;             APPLICATION NUMBER: 60/065,009
;             FILING DATE: 1997-11-14
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Keeley K. Brewer
;             REGISTRATION NUMBER: 40,932
;             REFERENCE/DOCKET NUMBER: PB359
```

```
;          TELECOMMUNICATION INFORMATION:
;               TELEPHONE: (301) 309-8504
;               TELEFAX: (301) 309-8512
;     INFORMATION FOR SEQ ID NO: 1:
;          SEQUENCE CHARACTERISTICS:
;               LENGTH: 2887 base pairs
;               TYPE: nucleic acid
;               STRANDEDNESS: double
;               TOPOLOGY: linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-070-927A-178

  Query Match             4.1%;  Score 58.4;  DB 2;  Length 2887;
  Best Local Similarity   50.0%;  Pred. No. 4.9;
  Matches   96;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy     106 IAAIIICCCIIICIGIACCCCAAAIAACCGGIICCIIIAAICGAAIACGAIAAIIAACGIIII 165
           :| |||   ||||   |   || |||| ||  |||| |      :  |  |  || | |
Db     533 IACIIIIIAIICCCIAACIIICCAAIAGCIIIIICCIAAAICCATTIATTTTTTTTIAIIAGC 592

Qy     166 CIAACAAIICCCICCCCAAAIICCCIAICAACAAICCCCAICCCCCCCCCCCAIIAIAI 225
           |  | | | || |  ||   |.   |     | | . |   ||  |||  . |  ||||
Db     593 AIAICIAIIIIICCAICCIIICCAICAICCIIACIAIAICCAIIACCTTTIAIIAAAAIAI 652

Qy     226 IIGAIAIAAIIGCIIIAAIAAIIGGIAIGAAAAGAIIIGGATTATICAIIICGIIGII 285
           :   | | |   ||| |||   .   .   :||  |   |       |  ||
Db     653 ICAGIAAIAGIAIAAGI.AAGAIGCCAACAAACCCAATICCCACAAICAICAAAAICAII 712

Qy     286 GGACGAACIGII 297
           | |   | | ||
Db     713 GCAGCAACIGII 724


RESULT 19
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
;  APPLICANT: SHIGENOBU, SHUJI
;  APPLICANT:  WATANABE, HIDEMI
;  APPLICANT:  BATTORI, MASAHIRA
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;  FILE REFERENCE: 081656/0259
;  CURRENT APPLICATION NUMBER: US/09/790,988
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: JP2000-047...
;  PRIOR FILING DATE: 2000-02-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

  Query Match             4.1%;  Score 58.4;  DB 1;  Length 640681;
  Best Local Similarity   45.0%;  Pred. No. 27;
  Matches  144;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy     167 AAIIC?CCIC IICIAIC , AAC:AA?C:::C:IAAICAAIAGAIAAIAAIIIIC 166
           |:| | | |  |||    |  |  |    | | |  |||  | ||   | | |
Db  142542 AC:IA:::A:A:C:A:::::::AA:A:::::A A:AAAAAAAAIAA?IA:GG?GAIAII 142581

Qy     167 IC?AAIC:C?C? C AAC::ACCAC?GA?AAIC:::::AC:::::::::?CAIAIAII 226
           |  ||:   |          | |  || |  :|  | |  | || || | || ||
Db  142582 IIICAA?AAAAI?AAAGACAA?AAICAGAGIAAIIIC?AAIATTCTGAIGIIAAIIII 142523

Qy     227 IGAIAIAAIIGIIIGIAAGAAIGIGIACIGAAAAGAIIIGGATTATICAIIICGIGIIG 286
           | | |  ||||    |  |  | | | |  ||  |      ||| | | ||
Db  142522 IIAIIGAIIGIIIIAAAAAIAIGAACIACAAC?AAAAAAAIIAATAAAIIIAAAACAA 142465

Qy     287 GACGAAIIGIIA?A:  :?A:::::A:::  :A:::::::::::::::::::?:?AA?GC 346
```

```
           ||: : |   | ||   ||  ||  | | |  . ||        |
Db  142462 C?AIIAIAIIA:::A:CAIA:AAAAAAAIA CC:AA:?AAA:?A?AIAIIAAIIAIIA 142405

Qy     347 AAIAI:AAAAA:::?A:C?IIC?IA:A:GA:GA:IAAAAAIIC?IC:::IC?AIIAIC?AII 406
           |   | ||| | |  |  |  || ||| ||   ||| |   | .
Db  142402 GAIC?AA?AAAAIC?IIAIC?IC?AAICAAC?IC?AAAGIIAIAIIIA?C?AIC?AAAA 142345

Qy     407 GAAICIAAC?A?IIC?AI 426
           | |   :  :|| |
Db  142342 IIAA?IC?C?AAGII?AAI 142325


Search completed: February 24, 2003, 18:17:35
Job time : 1744 secs
```